

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 03:49:46 ; Search time 1820.86 Seconds
(without alignments)
15998.962 Million cell updates/sec

Title: US-09-981-353-24
Perfect score: 1001
Sequence: 1 atgattacgaattcagctc.....gtgagcaaaaggccagcaaa 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sv.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	582	58.1	632	9	BC029149	BC029149 Homo sapi
2	545.8	54.5	560	6	AR182207	AR182207 Sequence
3	545.8	54.5	560	6	I95747	I95747 Sequence
4	397.6	39.7	105485	9	HUAC002301	AC02301 Homo sapi
5	387.4	38.7	174007	9	AC023831	AC023831 Homo sapi
6	387.4	38.7	175599	2	AC009133	AC009133 Homo sapi
7	368.8	36.8	673	10	BC031800	BC031800 Mus muscu
8	357.2	35.7	3409	12	CVU94951	U94951 Cloning vec
9	357.2	35.7	3595	12	AF174493	AF174493 Cloning v
10	357.2	35.7	3733	12	AF146395	AF146395 Cloning v
11	357.2	35.7	4630	12	CVU35316	U35316 Cloning vec
12	357.2	35.7	4630	12	CVU94950	U94950 Cloning vec
13	355.8	35.5	5086	6	AX107161	AX107161 Sequence
14	355.8	35.5	5086	6	AF438204	AF438204 Positive
15	352.6	35.2	3417	12	AF153422	AF153422 Cloning v
16	352.6	35.2	3417	12	AF153422	AF153422 Cloning v
17	352.4	35.2	5925	6	AX05712	AX05712 Sequence
18	352.4	35.2	5925	6	AX05712	AX05712 Sequence
19	352.4	35.2	6680	6	AX010972	AX010972 Sequence
20	352.2	35.2	3789	6	AX287967	AX287967 Sequence
21	352	35.2	3789	6	AR130382	AR130382 Sequence
22	352	35.2	3950	12	EVU81273	U81273 Expression
23	352	35.2	3966	12	EVU81274	U81274 Expression
24	352	35.2	3970	12	EVU81275	U81275 Expression
25	351.8	35.1	8710	6	AX195438	AX195438 Sequence
26	351.8	35.1	9019	6	I60241	I60241 Sequence 3
27	351.6	35.1	4190	12	XXU13848	U13848 pEXCell cto
28	351.4	35.1	4591	12	SYNPEZ218A	I60242 Sequence 4
29	351.4	35.1	8082	6	AX114871	AX114871 Sequence
30	351.4	35.1	8153	6	AX114871	AX114871 Sequence
31	351.2	35.1	2939	12	SYNPLINKY	D50400 Phagemid pL
32	351.2	35.1	3837	12	SYNPPDSZ	D50401 Phagemid pP
33	351	35.1	8800	6	AX469663	AX469663 Sequence
34	351	35.1	9208	12	AF064067	AF064067 Expressio
35	351	35.1	9211	12	AF064066	AF064066 Expressio
36	351	35.1	9215	12	AF049063	AF049063 Expressio
37	351	35.1	9556	12	AF049064	AF049064 Expressio
38	351	35.1	10042	6	AX469664	AX469664 Sequence
39	350.8	35.0	584	6	AX063012	AX063012 Sequence
40	350.8	35.0	584	6	AX367929	AX367929 Sequence
41	350.8	35.0	655	6	AX351379	AX351379 Sequence
42	350.8	35.0	738	6	AR036905	AR036905 Sequence
43	350.8	35.0	738	6	AR141144	AR141144 Sequence
44	350.8	35.0	738	6	AR181919	AR181919 Sequence
45	350.8	35.0	742	6	AX197652	AX197652 Sequence

ALIGNMENTS

RESULT 1
BC029149
LOCUS BC029149
DEFINITION Homo sapiens, similar to Homolog of rat Zymogen granule membrane protein, clone MGC:34820 IMAGE:5187790, mRNA, complete cds.
ACCESSION BC029149
VERSION BC029149.1
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS Strausberg, R.
TITLE Direct Submission

JOURNAL

Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 50 Row: h Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES

source Location/Qualifiers

1. .632

/organism="Homo sapiens"

/db_xref="LocusID:123887"

/db_xref="taxon:9606"

/clone="MGC:34820 IMAGE:5187790"

/tissue_type="Colon, Kidney, Stomach, adult, whole pooled"

/clone_lib="NIH_MGC_116"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

7. .510

/codon_start=1

/product="similar to Homolog of rat zymogen granule

membrane protein"

/protein_id="AAH29149.1"

/db_xref="GI:20810121"

/translation="MTVALLALLCASGNATQARSSSYSGYGGGKRFPSHSNQ

LGPIPALRVNVTYIVGLQVRYKWSYDVGRRNGDLEEIFLHPGESVIQVSGRYK

WYLKLVFTDKGRYLSFGKDSGTSFNAVPLHPNTVLRITSRSGSLDAIGLHWYV

PTSCSRC"

BASE COUNT 157 a 165 c 161 g 149 t

Query Match

Best Local Similarity 58.1%; Score 582; DB 9; Length 632;

Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 CCAGAGTTGACAGTCGCTCTCCAGCCCTTCTCTGTGCTCAGCCCTCGCAATGCC 87

DB 1 CCAGAGTTGACAGTCGCTCTCCAGCCCTTCTCTGTGCTCAGCCCTCGCAATGCC 60

QY 88 ATTCAGGCCAGGTCCTCTCTATAGTGAGAGTATGGAAGTGGTGGTGAAGCCGATTC 147

DB 61 ATTCAGGCCAGGTCCTCTCTATAGTGAGAGTATGGAAGTGGTGGTGAAGCCGATTC 120

QY 148 TCTCATCTTGGCAACAGTTGGACCGGCCCATCACCGCCCTCCGGGTCCAGGTCAACACA 207

DB 121 TCTCATCTTGGCAACAGTTGGACCGGCCCATCACCGCCCTCCGGGTCCAGGTCAACACA 180

QY 208 TACTACATCTAGTCTTACAGTGGCGCTATGGCAAGTGGGACGACATATGTTGGGTGGT 267

DB 181 TACTACATCTAGTCTTACAGTGGCGCTATGGCAAGTGGGACGACATATGTTGGGTGGT 240

QY 268 CGCAACGGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTGATCCAGGTTTCT 327

DB 241 CGCAACGGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTGATCCAGGTTTCT 300

QY 328 GGGAGTACAAAGTGGTACCTGAAGAAGCTGGTATTTGTGACAGAACAGGCCCGCTATCTG 387

DB 301 GGGAGTACAAAGTGGTACCTGAAGAAGCTGGTATTTGTGACAGAACAGGCCCGCTATCTG 360

QY 388 TCCTTTTGGGAAAGACAGTGGCACAAGTTTCAATCCCGTCCCTTGGCACCCCAACACCGTG 447

DB 361 TCCTTTTGGGAAAGACAGTGGCACAAGTTTCAATCCCGTCCCTTGGCACCCCAACACCGTG 420

QY 448 CTCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCCATTGGCCTGCACTGGGAT 507

DB 421 CTCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCCATTGGCCTGCACTGGGAT 480

QY 508 GTTTACCCCTAGCTGCAGCAGATGCTGAGCCCTCTCTCTCTGGCAGGGGCACTGTGAT 567

DB 481 GTTTACCCCTAGCTGCAGCAGATGCTGAGCCCTCTCTCTCTGGCAGGGGCACTGTGAT 540

QY 568 GAGGAGTAAGAATCCCTTATCACTAACCCCATCCAAATGG 609

DB 541 GAGGAGTAAGAATCCCTTATCACTAACCCCATCCAAATGG 582

RESULT 2

AR182207

LOCUS

Sequence 10 from patent US 6337195.

AR182207

ACCESSION

AR182207

VERSION

AR182207.1

GI:20225123

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 560)

AUTHORS

Yu, G.-L. and Rosen, C.

TITLE

Colon specific genes and proteins

JOURNAL

Patent: US 6337195-A 10 08-JAN-2002;

FEATURES

Location/Qualifiers

1. .560

source

/organism="unknown"

BASE COUNT 113 a 157 c 152 g 138 t

ORIGIN

Query Match

54.5%; Score 545.8; DB 6; Length 560;

Best Local Similarity 99.5%; Pred. No. 1.6e-148;

Matches 558; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 42 AGTCGCTCTCTAGCCCTTCTCTGTGCTCAGCCCTCTGGCAATGCCATTTCAGGCCAGTGC 101

DB 1 AGTCGCTCTCTAGCCCTTCTCTGTGCTCAGCCCTCTGGCAATGCCATTTCAGGCCAGTGC 60

QY 102 TTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGGAAAGCGATTCTCTCATTTCTGGCAA 161

DB 61 TTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGGAAAGCGATTCTCTCATTTCTGGCAA 120

QY 162 CCAGTTGGAGCGGCCCATCACCGCCCTCCGGGTCCGAGTCAACACATACATACATCGTAGG 221

DB 121 CCAGTTGGAGCGGCCCATCACCGCCCTCCGGGTCCGAGTCAACACATACATACATCGTAGG 180

QY 222 TCCTTCAGTTCGCTATGGCAAGTGTGGAGCGACTATGTGGTGGTGGTGGTGGTGGTGGTGG 281

DB 181 TCCTTCAGTTCGCTATGGCAAGTGTGGAGCGACTATGTGGTGGTGGTGGTGGTGGTGGTGG 240

QY 282 GGAGGAGATCTTCTTCACCCCTGGGGGAAATCAGTGATCCAGGTTTCTGGGGAAGTACAAGT 341

DB 241 GGAGGAGATCTTCTTCACCCCTGGGGGAAATCAGTGATCCAGGTTTCTGGGGAAGTACAAGT 300

QY 342 GTACCTCAACAAGCTGTGTTTGGACAGACAGGCCGCTATCTGTCTTTTGGGAAAGA 401

DB 301 GTACCTCAACAAGCTGTGTTTGGACAGACAGGCCGCTATCTGTCTTTTGGGAAAGA 360

QY 402 CAGTGGCACAAAGTTTCAATCCCGTCCCTTTCACCCCAACACCGTGTCTCCGCTTCATCAG 461

DB 361 CAGTGGCACAAAGTTTCAATCCCGTCCCTTTCACCCCAACACCGTGTCTCCGCTTCATCAG 420

QY 462 TGGCGGCTGTGTTCTCTCATGATGCCATTGGCCTGCATGGGATGTTTACCCCATAG 521

```
|||||
421 TGCCGGTGGTCTCTCATCGATGCCATTGGCCCTGCACCTGGATGTTTACCCCACTAG 480
|||||
522 CTGCACAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGAGTAAGAAT 581
|||||
481 CTGCACAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGAGTAAGAAT 540
|||||
582 CCCTTATCACTAACCCCACTC 602
|||||
541 -CCTTATCACTAACCCCACTC 560

RESULT 3
195747
LOCUS
DEFINITION
Sequence 10 from patent US 5733748.
ACCESSION
195747.1 GI:3940217
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 560)
AUTHORS
Yu, G.-L. and Rosen, C.
TITLE
Colon specific genes and proteins
JOURNAL
Patent: US 5733748-A 10 31-MAR-1998;
FEATURES
Location/Qualifiers
source
1..560
/organism="unknown"
BASE COUNT 113 a 157 c 152 g 138 t
ORIGIN
Query Match 54.5%; Score 545.8; DB 6; Length 560;
Best Local Similarity 99.5%; Pred. No. 1.6e-148;
Matches 558; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 42 AGTCGCTCTCTAGCCCTTCTCTGTGCTCAGCCTCTGCGCAATGCCATTCAGGCCAGGTC 101
|||||
DB 1 AGTCGCTCTCTAGCCCTTCTCTGTGCTCAGCCTCTGCGCAATGCCATTCAGGCCAGGTC 60
|||||
OY 102 TTCTCTCTATAGTGGAGATGATGAAGTGGTGTGGAAGCGATCTCTCATCTCTGCGAA 161
|||||
DB 61 TTCTCTCTATAGTGGAGATGATGAAGTGGTGTGGAAGCGATCTCTCATCTCTGCGAA 120
|||||
OY 162 CCAAGTTGGAGCGCCCATACCCCTCCGGGTCGAGTCAACACATACATACATCTAGG 221
|||||
DB 121 CCAAGTTGGAGCGCCCATACCCCTCCGGGTCGAGTCAACACATACATACATCTAGG 180
|||||
OY 222 TCTTCAGGTGCGCTATGGCAAGGTGTGGAGCGACTATGTGGGTGGTCGCAACGGAGACCT 281
|||||
DB 181 TCTTCAGGTGCGCTATGGCAAGGTGTGGAGCGACTATGTGGGTGGTCGCAACGGAGACCT 240
|||||
OY 282 GGAGGAGATCTTCTGCACCTGGGGAATCAGTATGATCCAGGTTTCTGGGAAGTACAAGTG 341
|||||
DB 241 GGAGGAGATCTTCTGCACCTGGGGAATCAGTATGATCCAGGTTTCTGGGAAGTACAAGTG 300
|||||
OY 342 GTACCTGAAGAAGCTGATTTGTGACAGAAAGGCGCGTATCTGCTTTTGGGAAGA 401
|||||
DB 301 GTACCTGAAGAAGCTGATTTGTGACAGAAAGGCGCGTATCTGCTTTTGGGAAGA 360
|||||
OY 402 CAGTGGCACAAGTTTCAATGCCGTCCCTTGTGACCCCAACACACCGTCTCCGCTTCATCAG 461
|||||
DB 361 CAGTGGCACAAGTTTCAATGCCGTCCCTTGTGACCCCAACACACCGTCTCCGCTTCATCAG 420
|||||
OY 462 TGCCCGGTCTGGTCTCTCATCGATGCCATTGGCCCTGCACCTGGATGTTTACCCCACTAG 521
|||||
DB 421 TGCCCGGTCTGGTCTCTCATCGATGCCATTGGCCCTGCACCTGGATGTTTACCCCACTAG 480
|||||
OY 522 CTGCACAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGAGTAAGAAT 581
|||||
DB 481 CTGCACAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGAGTAAGAAT 540
|||||
OY 582 CCCTTATCACTAACCCCACTC 602
|||||
```

```
|||||
541 -CCTTATCACTAACCCCACTC 560
|||||
RESULT 4
HUAC002301
LOCUS
DEFINITION
Homo sapiens Chromosome 16 BAC clone CIT987SK-A-328A3, complete
sequence.
ACCESSION
AC002301
VERSION
AC002301.1 GI:2979567
KEYWORDS
HTG.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 105485)
AUTHORS
Loftus, B.J., Kim, U.J., Sneddon, V.P., Kalush, F., Brandon, R.,
Fuhrmann, J., Mason, T., Crosby, M.L., Barnstead, M., Cronin, L.,
Deslattes Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S.,
Eichler, E.E., Harris, P.C., Venter, J.C. and Adams, M.D.
TITLE
Genome duplications and other features in 12 Mb of DNA sequence
from human chromosome 16p and 16q
JOURNAL
Genomics 60 (3), 295-308 (1999)
MEDLINE
99425270
PUBMED
10493829
REFERENCE
2 (bases 1 to 105485)
AUTHORS
Adams, M.D., Loftus, B.J., Zhou, L., Crosby, M., Fuhrmann, J.,
Mason, T.M., Brandon, R., Kim, U.J., Kerlavage, A.R. and Venter, J.C.
TITLE
Homo sapiens Chromosome 16 BAC clone CIT987SK-A-328A3
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 105485)
AUTHORS
Adams, M.D. and Loftus, B.J.
TITLE
Direct Submission
JOURNAL
Submitted (19-JUN-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, Email:
mdadams@tigr.org
REFERENCE
4 (bases 1 to 105485)
AUTHORS
Adams, M.D. and Loftus, B.J.
TITLE
Direct Submission
JOURNAL
Submitted (20-MAR-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
REFERENCE
5 (bases 1 to 105485)
AUTHORS
Adams, M.D.
TITLE
Direct Submission
JOURNAL
Submitted (31-MAR-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
COMMENT
On Mar 20, 1998 this sequence version replaced gi:2951943.
Address all correspondence to: Mark Adams The Institute for Genomic
Research 9712 Medical Center Dr., Rockville, MD 20850, USA e-mail
address: mdadams@tigr.org. The orientation of the sequence is
from SP6 end to T7 end. Genes were identified by a combination of
five methods including: XGRAIL (available by anonymous ftp from
arthur.epm.ornl.gov), Genefinder (Chris Burge,
Washington), Genscan (Chris Burge,
http://gnomic.stanford.edu/~chris/GENSCAN.html) searches of the
complete sequence against a peptide database, and the Human gene
index database at TIGR (http://www.tigr.org/db/hgi.html).
Genes without peptide homology having spliced EST hits are termed
'unknown gene product'. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Location/Qualifiers
1..105485
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p11.2"
/clone="A-328A3"
3032..3149
/notes="7608, STS1-CSRL-24g1-ua/CSRL-24g1-uz, Chr. 16, Homo
sapiens"
/db_xref="dbSTS:G02122"
14885..15002

STS
STS
```

```
/note="7608, STS1-cSRL-24g1-uA/cSRL-24g1-u2, Chr. -, Homo sapiens"
/db_xref="dbSTS:G02122"
75574..75694
/note="22506, SHGC-5801, Chr. 17, Homo sapiens"
/db_xref="dbSTS:G14087"
<77884..79132
/gene="A-328A3.1"
join(<77884..77938,78122..78254,78721..79132)
/gene="A-328A3.1"
join(77884..77938,78122..78254,78721..79036)
/gene="A-328A3.1"
/codon_start=1
/product="Homolog of rat zymogen granule membrane protein"
/protein_id="AAC08708.1"
/db_xref="GI:2979568"
/translocation="MLTVALLALICASASNAIQARSSYSGEYSGGGRFHSNGQ
LDGPITALRVNTYIVGLOVYKVMDSYVGGNRDLEEIFLHPGESVIOVSGKYK
WYKLKLVFVYDKRYLFGKDSGTSFNAVPLHPNTVLRFTSGRSGSLIDAIGLHWYV
PTSCSRC"
7451..87568
/note="7608, STS1-cSRL-24g1-uA/cSRL-24g1-u2, Chr. -, Homo sapiens"
/db_xref="dbSTS:G02122"
89374..>97443
/gene="A-328A3.2"
/note="(partial)"
join(89374..89467,95545..95740,97022..97149,97241..>97443)
/gene="A-328A3.2"
join(89398..89467,95545..95740,97022..97149,97241..97443)
/gene="A-328A3.2"
/codon_start=1
/product="Similar to kinesin-like DNA binding protein (KID)"
/protein_id="AAC08709.1"
/db_xref="GI:2979569"
/translocation="MAAGSTQQRREMAAASAAISGACRCLSKIGATRRPPARV
RVAVRPFVDGRLAGDPPPCVMDSCSLEIANRNHQTLYQFDFYGERSTQOD
IVAGSVOPILRHLEGONASVLAAYGPTGACKTHTMLGSPQPGVIPRALMDLLQTLRE
EGAGRPWALSVMYSLEYIEQVRPRAGWERNKSGK"
102130..104031
/gene="A-328A3.5"
join(102130..102255,102645..102712,103238..104031)
/gene="A-328A3.5"
join(102130..102255,102645..102712,103238..103673,
103910..104031)
/gene="A-328A3.5"
/note="102132"
join(102130..102255,102645..102712,103461..103673,
103762..103821,103910..104031)
/gene="A-328A3.5"
join(102132..102255,102645..102712,103461..103673,
103762..103821,103910..103957)
/gene="A-328A3.5"
/codon_start=1
/product="orip binding protein like gene product"
/protein_id="AAC08711.1"
/db_xref="GI:2979571"
/translocation="MLAQAQAEKENHCPTMLRPLSHRTVTRAKPLKAVVPLQLOE
QAASPNAEIHLNKGKRLKSLDLAPEKAEDWELQISPELLAHGKQLDLN
EGSARDLSLQRLGPKAQLIVGWRELHPFQSDLEVEGITGKQMESFLKANILG
LAAGQRCGAS"
102132..102255,102645..102712,103238..103264)
/gene="A-328A3.5"
/codon_start=1
/product="orip binding protein like gene product"
/protein_id="AAC08712.1"
/db_xref="GI:2979572"
/translocation="MLAQAQAEKENHCPTMLRPLSHRTVTRAKPLKAVVPLQLOE
QAASPNAEIHLNKGKRLKGANITGY"
104673..104723
/note="50 N's represent a gap of approximately 800 bp"
<105069..>105276
```

```
/gene="A-328A3.4"
/note="(partial)"
<105069..>105276
/gene="A-328A3.4"
105078..>105276
/gene="A-328A3.4"
/codon_start=1
/product="Similar to fragment of MAZ DNA binding protein"
/protein_id="AAC08710.1"
/db_xref="GI:2979570"
/translocation="MCGAPRDVYHLNRHLKLSHDEKPYQCPVCQQRKFRKDRMSYHV
RSHDGAHVHKYPNCSHCKGFSR"
BASE COUNT 27103 a 25382 c 25633 g 27316 t 51 others
ORIGIN

Query Match 39.78; Score 397.6; DB 9; Length 105485;
Best Local Similarity 88.56; Pred. No. 7.9e-105;
Matches 406; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 213 CATCGTAGGTCTTCAGGTGCGCTATGCCAAGGTGTGGAGCGACTATGTGGTGGTTCGCAA 272
Db 78712 CTTCTCAGTCTTCAGGTGCGCTATGCCAAGGTGTGGAGCGACTATGTGGTGGTTCGCAA 78771

Qy 273 CGGAGACCTGGAGGAGATCTTTCTGCACCTGGGGAATCAGTATCCAGGTTTCTGGGAA 332
Db 78772 CGGAGACCTGGAGGAGATCTTTCTGCACCTGGGGAATCAGTATCCAGGTTTCTGGGAA 78831

Qy 333 GTACAAGTGTGACCTGAAGAGCTGGTATTTGTACACAGCAAGGCGCGCTATCTCTCTT 392
Db 78832 GTACAAGTGTGACCTGAAGAGCTGGTATTTGTACACAGCAAGGCGCGCTATCTCTCTT 78891

Qy 393 TGGGAAAGACAGTGGCACAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGTCTCCG 452
Db 78892 TGGGAAAGACAGTGGCACAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGTCTCCG 78951

Qy 453 CTTTCATCAGTGGCGGTCTGGTTCTCTCATCGATGCCATTGGCCTGCATCGGATGTTTA 512
Db 78952 CTTTCATCAGTGGCGGTCTGGTTCTCTCATCGATGCCATTGGCCTGCATCGGATGTTTA 79011

Qy 513 CCCCACTAGCTGCAGCAGATCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 572
Db 79012 CCCCACTAGCTGCAGCAGATCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 79071

Qy 573 GTAAGAACTCCCTTATCATAACCCCAATCAAAATGNNNNNNNNNNNNNNNNNNNNNNNN 632
Db 79072 GTAAGAACTCCCTTATCATAACCCCAATCAAAATGNNNNNNNNNNNNNNNNNNNNNNNN 79131

Qy 633 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 671
Db 79132 GCTAGTCTGTGTGGGGCATCTGTGGTGGGATATCTGC 79170

RESULT 5
AC023831/c AC023831 174007 bp DNA linear PRI 23-AUG-2002
LOCUS Homo sapiens chromosome 16 clone RP11-74E23, complete sequence.
DEFINITION AC023831
ACCESSION AC023831
VERSION AC023831.9 GI:22450570
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 174007)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174007)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 174007)
```


QY 393 TGGGAAAGACAGTGGCACAAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGTGCTCCG 452
 |||||
 Db 90431 TGGGAAAGACAGTGGCACAAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGTGCTCCG 90490
 |||||
 QY 453 CTTTCATCAGTGGCGCGTGTGTTCTCTCATCGATGCCATTTGGCTGCAGTGGGATGTTTA 512
 |||||
 Db 90491 CTTTCATCAGTGGCGCGTGTGTTCTCTCATCGATGCCATTTGGCTGCAGTGGGATGTTTA 90550
 |||||
 QY 513 CCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTGCGAGGCGCACTGTGATGAGGA 572
 |||||
 Db 90551 CCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTGCGAGGCGCACTGTGATGAGGA 90610
 |||||
 QY 573 GTAAGAACTCCCTTATCACTAACCCCAATGGS 609
 |||||
 Db 90611 GTAAGAACTCCCTTATCACTAACCCCAATGGS 90647
 |||||

RESULT 7
 BC031800
 LOCUS
 DEFINITION Mus musculus, RIKEN cDNA 1810010M01 gene, clone MGC:19113
 IMAGE:4208847, mRNA, complete cds.
 ACCESSION BC031800
 VERSION BC031800.1 GI:21594153
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

BC031800 673 bp mRNA linear ROD 07-AUG-2002
 Mus musculus, RIKEN cDNA 1810010M01 gene, clone MGC:19113
 IMAGE:4208847, mRNA, complete cds.
 ACCESSION BC031800
 VERSION BC031800.1 GI:21594153
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabp3r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: angbcm@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 24 Row: P Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES

Source

1...673
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:19113 IMAGE:4208847"
 /tissue_type="Colon, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP Co24"
 /lab_host="DH105"
 /note="Vector: pCMV-SPORT6"
 5...508
 /codon_start=1
 /product="RIKEN cDNA 1810010M01 gene"
 /protein_id="AAH31800.1"
 /db_xref="GI:21594154"

CDS

/db_xref="LocusID:69036"
 /translation="MLAVALLVLLCASANSIQSRTSSYSGBYGKGRKRFHSGNQ
 LDCPIAFRRVRNRYTVGLQVRYGTWSDYVGTDGLDEIFLHPGESYIOVSGKYK
 SYVKQMFVTDKGRYLPFGKASGTSFNAPVLPNTVLRFLISGRSGSAIDISLHWDY
 PSHCNRC"

BASE COUNT 185 a 165 c 162 g 161 t
 ORIGIN
 Query Match 36.8% Score 368.8; DB 10; Length 673;
 Best Local Similarity 82.9%; Pred. No. 1.1e-96;
 Matches 421; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
 QY 30 CAGAATGTTGACAGTGCCTCTCCTAGCCCTCTCTGTGCTCAGCCCTTGGCAATGCCAT 89
 |||||
 Db 1 CAGAATGTTGGCGGTGCGCCCTCTTAGTCTCTTTGTGCTCGCCCTCTGCTATTCAT 60
 |||||
 QY 90 TCAGGCCAGTCTTCCTCCTATAGTGCAGAGTATGGAAGTGGTGGTGAAGCGCATTC 149
 |||||
 Db 61 TCAGTCCAGGACATCTCTTACAGTGGAGAGTATGGAGGTAAAGGAGGAAGCGCATTC 120
 |||||
 QY 150 TCATTTCTGGCAACAGTTGGACGCCCATCACCGCCCTCCGGGTCCGAGTCAACACATA 209
 |||||
 Db 121 TCATTTCTGGCAACAGTGGACGCCCATCACCGCCCTTCGGTATCCGGGTCAACACATA 180
 |||||
 QY 210 CTACATGTTAGTCTTTCAGTGGCTATGCAAGTGTGGAGGACTATGTGGGTGTCG 269
 |||||
 Db 181 CTACATGTTAGTCTTCCAGTGGCTATGCAAGTGTGGAGTCTATGTAGGTGGCAC 240
 |||||
 QY 270 CAACGGAGACTGGAGGAGATCTTCTGCACCCCTGGGGAATCAGTATCCAGTTCTGG 329
 |||||
 Db 241 ACAGGGAGACTGGAGGAGATCTTCTGCACCCCGGGGAATCTGTGATCAGGTGCTGG 300
 |||||
 QY 330 GAAGTACAAGTGTGCTGTAAGAGTGGTATTGTGACAGACAAGGCCGCTATCTGTG 389
 |||||
 Db 301 CAAGTACAATCATACGTGAAGCAGATGATCTCTGTCAGCAGACAAGGCCGCTACCTGCC 360
 |||||
 QY 390 TTTTGGAAAGACAGTGGGACAAAGTTTCANAGCGGTCCCTTGGACCCCAACACCGTGT 449
 |||||
 Db 361 TTTTGGAAAGACAGTGGGACAAAGTTTCAACGGTGTTCCTTGCATCCCAACACTGTCT 420
 |||||
 QY 450 CCGCTTCATCAGTGGCGGTCTGTTCTCTCATCGATGCCATTTGGCTGCAGTGGGATGT 509
 |||||
 Db 421 CCGTTTCATTTAGTGGGCGATCTGGCTCGCATAGATTCTATTAGCCTGCATGGGATAC 480
 |||||
 QY 510 TTATCCCACTAGCTGCAGCAGATGCTGA 537
 |||||
 Db 481 CTACCTGACTGCAACACTTGTCTGA 508
 |||||

RESULT 8
 CVU94951/c
 LOCUS
 DEFINITION Cloning vector pZip1, complete sequence.
 ACCESSION U94951
 VERSION U94951.1 GI:2218120
 KEYWORDS synthetic construct.
 SOURCE synthetic construct
 ORGANISM artificial sequences.

REFERENCE
 AUTHORS Kerschbaumer, R.J., Hirschl, S., Kaufmann, A., Ibl, M., Koenig, R. and
 Himmler, G.
 TITLE Single-chain Fv fusion proteins suitable as coating and detecting
 reagents in a double antibody sandwich enzyme-linked immunosorbent
 assay
 JOURNAL Anal. Biochem. 249 (2), 219-227 (1997)
 MEDLINE 97356438
 PUBMED 9212874
 REFERENCE 2 (bases 1 to 3409)
 AUTHORS Kerschbaumer, R.J. and Himmler, G.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAR-1997) Strain Improvement, Institute of Applied
 Microbiology, Muthgasse 18, Vienna A-1190, Austria

FEATURES
Source

Location/Qualifiers
1. .3409
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="designed to fuse single-chain Fv to a hinge-region and a leucine-zipper"
201..1061
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC33185.1"
/db_xref="GI:2218121"
/translation="MSIQHFRVALIPFAAFCLVPFAHPETLVKVKDAFDOLGARGVY IELDNSKILLESFRPEPRPMSTFKVLICGAVLSRIDAGQQLGRRIRHYSDNLVE YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGPKELTAFLNMGDHVTRL DRWPELNEALPNDERTTPVAMATTLKLTGELLTASRQQLIDWADKVGAPL LRSALPAGNFIADKSGAGERSGIIAALPGDKPSRIIVYITTSQATMDERNQIA EIGASLIKHW"
2274..2528
/note="pelB-leader is followed by a polylinker"
/codon_start=1
/transl_table=11
/product="hinge-region and leucine-zipper with N-terminal PelB-leader and C-terminal hexahistidine tag"
/protein_id="AAC33186.1"
/db_xref="GI:2218122"
/translation="WKYLLPTAAAGLLLLAQAQVQLQARLAAAPKPTPPGSSR MKQLEDKVEELSKNHYHLENEVARLKLIVGRGGHHHHH"
830 a 880 c 840 g 859 t

CDS

CDS

BASE COUNT
ORIGIN

Query Match 35.7%; Score 357.2; DB 12; Length 3409;
Best Local Similarity 84.8%; Pred. No. 3.4e-93;
Matches 374; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 561 CTGTGATGAGGAGTAAAGTCTTATCACTACCCCAATCCCAATGNNNNNNNN 620
DB 2298 CTGCGCTAGGCAATAGTATTTCATGACTGCTCTTCAATAGAAATTTTCATGGAA 2239
QY 621 NNN 2238
DB 2238 GCTTGGCGTAATCATGTCATAGCTGTTCTGTGTGAAATTTTATCGCTCACAAATTC 680
QY 681 CACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCT 740
DB 2178 CACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCT 740
QY 741 AACTCACATTAATGCTTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2119
DB 2118 AACTCACATTAATGCTTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
QY 801 AGCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATGGCGCTCTT 2059
DB 2058 AGCTGCATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTGGGTATGGCGCTCTT 860
QY 861 CCCTTCTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999
DB 1998 CCCTTCTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
QY 921 CTCACATCAAGGGGTAATACGTTATCCAGAGATCAGGGGATACCCAGAGAGACA 1939
DB 1938 CTCACATCAAGGGGTAATACGTTATCCAGAGATCAGGGGATACCCAGAGAGACA 980
QY 981 TGTGAGCAAAAGCCAGCAAA 1001
DB 1878 TGTGAGCAAAAGCCAGCAAA 1858

RESULT 9
AF174493/c
LOCUS
DEFINITION
ACCESSION

VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORSTITLE
JOURNALFEATURES
source

CDS

CDS

BASE COUNT
ORIGINQuery Match
Best Local Similarity
MatchesScore 357.2; DB 12; Length 3595;
Pred. No. 3.5e-93;
Conservative 0; Mismatches 67; Indels 0; Gaps 0;QY 561 CTGTGATGAGGAGTAAAGTCTTATCACTACCCCAATCCCAATGNNNNNNNN 620
DB 2298 CTGCGCTAGGCAATAGTATTTCATGACTGCTCTTCAATAGAAATTTTCATGGAA 2239
QY 621 NNN 2238
DB 2238 GCTTGGCGTAATCATGTCATAGCTGTTCTGTGTGAAATTTTATCGCTCACAAATTC 680
QY 681 CACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCT 740
DB 2178 CACACACATACGAGCGGAGCAATAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCT 2119
QY 741 AACTCACATTAATGCTTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800
DB 2118 AACTCACATTAATGCTTCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059

```

Qy 801 AGTGCATTAAATGAATGGCCAAACGCGCGGGGAGAGCGGCTTTGCGTATTGGCGCTCTT 860
Db 2058 AGTGCATTAAATGAATGGCCAAACGCGCGGGGAGAGCGGCTTTGCGTATTGGCGCTCTT 1999
Qy 861 CCCTTCTCGCTCACTGACTCGCTGCGCTGCGTCTGCTGCGGCGGAGCGGTATCAG 920
Db 1998 CCCTTCTCGCTCACTGACTCGCTGCGCTGCGTCTGCTGCGGCGGAGCGGTATCAG 1939
Qy 921 CTCACCTCAAAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGGGATACGCAGGAAAGAACA 980
Db 1938 CTCACCTCAAAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGGGATACGCAGGAAAGAACA 1879
Qy 981 TGTGAGCAAAAGGCCAGCAAA 1001
Db 1878 TGTGAGCAAAAGGCCAGCAAA 1858

RESULT 10
AF146395/c
LOCUS 3733 bp DNA circular SYN 08-JUN-1999
DEFINITION Cloning vector pCLZip complete sequence.
ACCESSION AF146395
VERSION AF146395.1 GI:5006904
KEYWORDS Cloning vector pCLZip.
SOURCE Cloning vector pCLZip.
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 3733)
AUTHORS Kerschbaumer, R.J., and Himmler, G.
TITLE Dedicated expression vectors for the production of diagnostic reagents
JOURNAL (in) Harper, K. and Ziegler, A. (Eds.);
RECOMBINANT ANTIBODIES: APPLICATIONS IN PLANT SCIENCE AND PLANT
PATHOLOGY: 57-82;
(1999)
2 (bases 1 to 3733)
AUTHORS Kerschbaumer, R.J., Hirschl, S. and Himmler, G.
TITLE Direct Submission
JOURNAL Submitted (28-APR-1999) Strain Improvement, Institute of Applied
Microbiology, Muthgasse 18, Vienna, AT A-1190, Austria
FEATURES
source 1..3733
/organism="Cloning vector pCLZip"
/db_xref="taxon:95174"
201..1061
/gene="bla"
201..1061
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAB37723.1"
/db_xref="GI:5006905"
/translation="MSIQHPRVALIPFFAFCPLVFAHPETLVKVKDAEDQLGARVGY
IEDLSKGLFSREFPRPMSTFKVLGGAVLRIDAGQQLGRRIHYSNDLVE
YSPVTKHTLDTMTAEELSAITMSDNTAALLLTIGGKELTAPFLHNMGDHVTSL
DRWPELNEAIPNDERDTPVAMATLRKLTLGLTILASRQOLIDWMEADKAVGPL
LRSLAPAGNFIADKSGAGERSGIITAAALPGDGPSPRVIVYITGSAQTMERNRQIA
EIGASLIKHW"
2274..2852
/notes="contains PelB leader region, constant domain from
human IgG kappa light chain, hinge-region, leucine-zipper
helical region and hexahistidine tag"
/codon_start=1
/transl_table=11
/product="PelB-IgG kappa light chain fusion protein"
/protein_id="AAB37724.1"
/db_xref="GI:5006906"
/translation="MYLLPTAAAGLLLLAAQAPMAQVQLQARLAATVAAPSVFIFPP
PSDQKSGTASVCLNNFYPREAKVQWVDNALQNSQESVTEODSKDSTYLSLSS
TFLSRADYKHKVYACEVTHOGLSSPVTKSFNRGSGGPKPTPPGSSRMKQLEDKV
EELLKSNHLENAVRLKLVGRGHHHHH"
917 a 974 c 924 g 918 t
BASE COUNT

```

```

ORIGIN
Query Match 35.7%; Score 357.2; DB 12; Length 3733;
Best Local Similarity 84.8%; Pred. No. 3.5e-93;
Matches 374; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 561 CTGTGATGAGGAGTAAGAACTCCCTTATCACTAAACCCCATCCAAATGGNNNNNNNN 620
Db 2298 CTGCGGTAGCAATAGTATTTCATTATGACTGTCTCTTCAAAATAGAAATTTTCATGAA 2239
Qy 621 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 680
Db 2238 GCTTGGCGTAAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTATCCCGTCAAAAT 2179
Qy 681 CACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGT 740
Db 2178 CACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGT 2119
Qy 741 AACTCATTAAATTTGCTGCTGCTCACTGCGCGCTTTCCAGTCCGGGAAACCTGTCTGTC 800
Db 2118 AACTCATTAAATTTGCTGCTGCTCACTGCGCGCTTTCCAGTCCGGGAAACCTGTCTGTC 2059
Qy 801 AGCTGCATTAAATGAATCGCCCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTT 860
Db 2058 AGCTGCATTAAATGAATCGCCCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTT 1999
Qy 861 CCGCTTCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGGCGAGCGGTATCAG 920
Db 1998 CCGCTTCTCGCTCACTGACTCGCTGCGTTCGCTGCGGCGAGCGGTATCAG 1939
Qy 921 CTCACCTCAAAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGGGATACGCAGGAAAGAACA 980
Db 1938 CTCACCTCAAAAGGGGGTAAATACGGTTATCCACAGAAATCAGGGGGATACGCAGGAAAGAACA 1879
Qy 981 TGTGAGCAAAAGGCCAGCAAA 1001
Db 1878 TGTGAGCAAAAGGCCAGCAAA 1858

RESULT 11
CVU35316/c
LOCUS 4630 bp DNA circular SYN 28-AUG-1998
DEFINITION Cloning vector pDAP2, complete sequence.
ACCESSION U35316
VERSION U35316.1 GI:1016008
KEYWORDS Cloning vector pDAP2.
SOURCE Cloning vector pDAP2.
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 4630)
AUTHORS Kerschbaumer, R.J., Hirschl, S., Schwager, C., Ibl, M. and Himmler, G.
TITLE pDAP2: a vector for construction of alkaline phosphatase fusion-proteins
JOURNAL Immunotechnology 2 (2), 145-150 (1996)
MEDLINE 98040853
PUBMED 9373322
REFERENCE 2 (bases 1 to 4630)
AUTHORS Kerschbaumer, R.J., Hirschl, S. and Himmler, G.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1995) Randolph J. Kerschbaumer, Strainimprovement,
Institute of Applied Microbiology, Nussdorfer Laende 11, Vienna, AT
A-1190, Austria
FEATURES
source 1..4630
/organism="Cloning vector pDAP2"
/db_xref="taxon:42878"
/plasmid="pDAP2"
/notes="vector for fusion of proteins or peptides to the
N-terminus of E.coli alkaline phosphatase"
201..1061
/codon_start=1
/transl_table=11
/product="beta-lactamase"
CDS

```

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	REFERENCE AUTHORS TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	FEATURES source
--	-------------------------------	--	--------------------

561
2298
621

[illegible]

U94350	
U94950.1	GI:2218117
synthetic construct ⁶	
synthetic construct	
artificial sequences	
1 (bases 1 to 4630)	
Kerschbaumer,R.J., Hirschl,S., Kaufmann,A., Ibl,M., Koenig,R. and Himmeler,G.	
Single-chain Fv fusion proteins suitable as coating and detecting reagents in a double antibody sandwich enzyme-linked immunosorbent assay	
Anal. Biochem.	249 (2), 219-227 (1997)
97356438	
9212874	
2 (bases 1 to 4630)	
Kerschbaumer,R.J. and Himmeler,G.	
Direct Submission	
Submitted (24-MAR-1997) StrainImprovement, Institute of Applied Microbiology, Muthgasse 18, Vienna A-1190, Austria	
1	
Location/Qualifiers	
1	

```

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="designed to fuse proteins to the N-terminus of a
modified E.coli alkaline phosphatase with an increased
specific activity"
201..1061

```

```

/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC33183.1"
/db_xref="GI:2218118"
/translation="MSIQHFRVALIPEFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IELDNLGSKILESFRRPEERPMSTTEKVLCCGAVLSRIDAGQELQRRITHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITLSDNTAANLLTIIGPKELATFLHMGCDHVTPL
DRWEPELNEAIPNDERDITMPVAMATLKRLTGELLTLASRQQLIDMNEADGHVTRL
LRGSLPAGWFIADKSGAGERSGRGIIAALGLPGDKPSRIVYITTSQATMDERNRQIA
EIGASLIKHW"
2274. .3749

```

```

/note="pelB-leader is followed by a polylinker; coding
sequence of E.coli JM81 alkaline phosphatase gene mutated
to change aspartate in the active site to serine to
enhance the specific activity of the enzyme; Mandecki et
al., (1991) Protein Eng. 4, 801-804"
/codon_start=1
/transl_table=11
/product="modified E.coli alkaline phosphatase with
N-terminal PelB-leader and C-terminal hexahistidine tag"
/protein_id="AAC33184.1"
/db_xref="GI:2218119"
/translation="MKYLLPTTAAAGLLLLAAQAPMAQVQLQARLAAARAPMPVLEN
RAAQGDITAPGGRRLTGDQTAALRSLSDKPKNITLLIGDNGDSEITAAARYAEG
AGGFFGDTALPGTQVTHYALKNCTKGPVYSSAAASATASWTGYKTNALGLGVDIH
EKDHPITLEMAKAAGALGNVSPKILQDAPLVAHVYTRSKCYGSPKCPGNAL
EKGKGSITTEOLLNARADYVLGGAKITFAETATGWEQKTLREQAARQYVLSDAA
SLNSVTBANKQPKLGLFADGNPVRVWLGPVKATYHGNIDKPAVCTPNPQRNDSVPTL
AMQTDKALELISNKEGFFQVQVEGASIDKDHAAANPCGQIGETVDLDEAVQALREFAK
KEGNTLVIVTADHAHASQIVAPATPKGTQALNTKDGAVMYNVSNGSDSEQHEHTGS
QLRTAYGPHAAANVGLTQDTLFTYTKAALGDIAHHHHH"

```

[illegible]

```
Db      2238  GCTGGCGTAATCATGTCATAGCTGTTTCTCTGTGTAATTTATTCGCGTCACAAATC 2179
QY      681  CACACAATACAGCGCGGAAGCATAAAGTGTAAAGCTCGGCTGCTAATGAGTGAGCT 740
Db      2178  CACACAATACAGCGCGGAAGCATAAAGTGTAAAGCTCGGCTGCTAATGAGTGAGCT 2119
QY      741  AACTCACATTAATGCGGTGCGCTACAGTCCGCCCTTTCCAGTCCGGAAACCTGTCGTGCC 800
Db      2118  AACTCACATTAATGCGGTGCGCTACAGTCCGCCCTTTCCAGTCCGGAAACCTGTCGTGCC 2059
QY      801  AGCTGCATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCTT 860
Db      2058  AGCTGCATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGCTATTGGCGGCTCTT 1999
QY      861  CCCTCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
Db      1998  CCCTCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
QY      921  CTCACCTAAAGCGCGTAATACGGTTATCCACAGAAATCAGCGGGGATACGCAAGGAAAGCA 980
Db      1938  CTCACCTAAAGCGCGTAATACGGTTATCCACAGAAATCAGCGGGGATACGCAAGGAAAGCA 1879
QY      981  TGTGACCAAAAGCCAGCAAA 1001
Db      1878  TGTGACCAAAAGCCAGCAAA 1858

RESULT 13
AX107161/c
LOCUS   AX107161 5086 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0125451.
ACCESSION AX107161
VERSION   AX107161.1 GI:13922646
KEYWORDS  synthetic construct.
SOURCE    synthetic construct.
ORGANISM  synthetic construct.
REFERENCE 1 (bases 1 to 5086)
AUTHORS  Hornby,D.P. and Matin,M.M.
TITLE     Cloning vectors
JOURNAL   Patent: WO 0125451-A 1 12-APR-2001;
          TRANSGENOMIC, Inc. (US)
FEATURES  Location/Qualifiers
           source
           1..5086
           /organism="synthetic construct"
           /db_xref="taxon:32630"
           /note="pMTet1"
BASE COUNT 1304 a 1275 c 1247 g 1260 t
ORIGIN

Query Match 35.5%; Score 355.8; DB 6; Length 5086;
Best Local Similarity 80.2%; Pred. No. 9.3e-93;
Matches 390; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY      516  CACTAGCTGCAGCAGATGTCAGCCCTCTCTCTCTGCGAGGCGCATGTGATCAGGAGTA 575
Db      3910  CACTAATCTCAATAATTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3851
QY      576  AGAATCCCTTATCACTAACCCTCCATCCAAATCGNNNNNNNNNNNNNNNNNNNNNNNN 635
Db      3850  ATTTCAATATTTTCAGGTTTCATGATGATCCATGTCGGCGGCGCACACAGCTTGGCGTAATCAT 3791
QY      636  NNNNNNNNNNNNTCTGTGTAATTTGTTATCCGCTCAAAATTCACACAAACATACGAG 695
Db      3790  GGTCATAGCTGTTTCTGCTGTGAATTTGTTATCCGCTCAAAATTCACACAAACATACGAG 3731
QY      696  CCGGAAGCATAAAGTGAAGCCCTGGGTCCTTAATGATGAGTCACTCACTCAATTAATG 755
Db      3730  CCGGAAGCATAAAGTGAAGCCCTGGGTCCTTAATGATGAGTCACTCACTCAATTAATG 3671
QY      756  CGTTGCGCTCACTGCGCGCTTTTCAGTCCGGAACCTGTCGTCGCCAGCTGCATTAATGAA 815
Db      3670  CGTTGCGCTCACTGCGCGCTTTTCAGTCCGGAACCTGTCGTCGCCAGCTGCATTAATGAA 3611
```

```
QY      816  TCGCCCAACGCGCGGAGAGCGGCTTTGCGTATTGGCGCGCTTTCGCTTCCTCGCTCA 875
Db      3610  TCGCCCAACGCGCGGAGAGCGGCTTTGCGTATTGGCGCGCTTTCGCTTCCTCGCTCA 3551
QY      876  CTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 935
Db      3550  CTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 3491
QY      936  TAATACCGTTATCCACAGAATCAGCGGATACGAGGAAAGCAATGTGAGCAAAAGGCC 995
Db      3490  TAATACCGTTATCCACAGAATCAGCGGATACGAGGAAAGCAATGTGAGCAAAAGGCC 3431
QY      996  AGCAAA 1001
Db      3430  AGCAAA 3425

RESULT 14
AF438204/c
LOCUS   AF438204 5086 bp DNA circular SYN 14-NOV-2001
DEFINITION Positive selection vector pMTet1, complete sequence.
ACCESSION AF438204
VERSION   AF438204.1 GI:16923909
KEYWORDS  Positive selection vector pMTet1.
SOURCE    Positive selection vector pMTet1.
ORGANISM  Positive selection vector pMTet1.
REFERENCE 1 (bases 1 to 5086)
AUTHORS  Matin,M.M. and Hornby,D.P.
TITLE     A positive selection vector combining tetracycline resistance that
          eliminates the need for bacterial plating
JOURNAL   Anal. Biochem. 278 (1), 46-51 (2000)
MEDLINE   20108508
PUBMED    10640352
REFERENCE 2 (bases 1 to 5086)
AUTHORS  Matin,M.M. and Hornby,D.P.
TITLE     Direct Submission
JOURNAL   Submitted (22-OCT-2001) Molecular Biology, Krebs Institute, Firth
          Court, Western Bank, Sheffield S10 2TN, UK
FEATURES  Location/Qualifiers
           source
           1..5086
           /organism="Positive selection vector pMTet1"
           /db_xref="taxon:177164"
           92..1282
           /gene="tetA(C)"
           92..1282
           /gene="tetA(C)"
           /note="modified form of the pBR322 tetA(C)"
           /codon_start=1
           /transl_table=11
           /product="tetracycline resistance protein"
           /protein_id="AAU31631.1"
           /db_xref="GI:16923910"
           /translation="MKSNNALIVLTGTVTDVGVIGLVMPVLPGLLRDVRHSDSIASH
           YGVLLALYALMQFLCAPLDRFRFRVLLASLLGATIDYAIMATTPLVLTLYAG
           RIVAGITGATGAVAGYAIADTDCEDRHRFLMSACFGVGVAGPVAGLGAISLH
           AFLAAALNGLNLLCGFLMGSKRGRRPMLRFPNVSFFRWAGMTVAAALMTV
           FFMOLVGQVPAALWTFEGEDRFRWSATMIGLSLAVFGILHAAQAFVTPATKRFGE
           KQIIAGMAADALGYVLLAFATRGWAPFIMILLASGGIGMALQAMLSRQVDDHOG
           QLOGSLAALSLTSITGLTVITAIYASASTWNLGAWLIVGALYLVCLPRLRGANWR
           ATST"
           3829..5085
           /note="from M. MspIm"
           /codon_start=1
           /transl_table=11
           /product="CS DNA methyltransferase"
           /protein_id="AAU31632.1"
           /db_xref="GI:16923911"
           /translation="MKPEILKLRISKLDTOKOASELIEVSKRAQQWESGKTEMHPA
           YYSFLOPKLKDKNFELSQAQTKLOKKIFDKYNQNTKNAEELAEITHIEERKDAYS
           SDFKFDLFSGIGIRQSFVNGKCFESSEIDPPAFKFTYTNFVGVVPGDITKVEAT
           TIFEHDLICAGFFGQPMWSHMKREGFEHPTQGTGNFHEIRLIIETKTKTVPVLENVPLG
```

INHDDGNTLKVIETLEDWGYKVHHTVLDASHFGIPQKRRELYLVAFLNQNINHFEPPK
PPMISKDGEVLESDVTGYSSEHLQSKYLFKKDDGKPSLIDKNTTGAVKTVLSTVHK
TQRLTGTVGKDEGATIRLLTNECKALMGPKPOFVIPVSRTOQMYRQMGNSVVVPVTK
1275 C 1248 G 1259 T

BASE COUNT
ORIGIN

	Query Match	35.5%;	Score 355.8;	DB 12;	Length 5086;
	Best Local Similarity 80.2%;	Pred. No. 9.3e-93;			
	Matches 390;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;
516	CAC	TACTAGCTGCAGAGATGCTGACGCTCTCTCTCTGCGAGGGGCACTGTGATGAGGAGTA	575		
3910	CAC	TAACTTCAATAATTTTCAGATGCTTCCTTTGAGTTAGATCCAAATTTACTACGAATCA	3851		
576	AGA	ACTTCCCTTATCACTAACCCCACTCCAAATGCGNNNNNNNNNNNNNNNNNNNNNNNN	635		
3850	ATT	TCAATATTTCAGGTTTCATGATGCTTCGCGCGGCACACGACGCTTGGCGTAATCAT	3791		
636	NNNNNNNNNN	NNTCCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACACATACGAG	695		
3790	GGT	CATAGCTCTTTCCCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACACATACGAG	3731		
696	CCG	AAGCAATAAAGCTGTAAAGCCTGGGGTCCCTTAATGAGTGAGCTAACTCACATTAATTG	755		
3730	CCG	AAGCAATAAAGCTGTAAAGCCTGGGGTCCCTTAATGAGTGAGCTAACTCACATTAATTG	3671		
756	CGT	TGCGCTCACATGCCCGGCTTTCCAGTCGGGAAACCTGTCGGGCGAGCTGCATTAAATGAA	815		
3670	CGT	TGCGCTCACATGCCCGGCTTTCCAGTCGGGAAACCTGTCGGGCGAGCTGCATTAAATGAA	3611		
816	TCG	GGCCACGCGCGGGGAGAGCGGTTTGGCGTATTGGCGCTCTTCGCGTCTTCCTCGCTCA	875		
3610	TCG	GGCCACGCGCGGGGAGAGCGGTTTGGCGTATTGGCGCTCTTCGCGTCTTCCTCGCTCA	3551		
876	CTG	ACTCGCTGCGCTCGCTCGCTTGGCTGCGGCGAGCGGTATCAGCTCACTCAAGCGCG	935		
3550	CTG	ACTCGCTGCGCTCGCTCGCTTGGCTGCGGCGAGCGGTATCAGCTCACTCAAGCGCG	3491		
936	TAAT	ACGGTTATCCACAGAAATCAGGGGGATACGACGAAAGACATGTGAGCAAAAGGCC	995		
3490	TAAT	ACGGTTATCCACAGAAATCAGGGGGATACCCAGAAAGACATGTGAGCAAAAGGCC	3431		
996	AGCAAA	1001			
3430	AGCAAA	3425			

[illegible]

SOURCE	Cloning vector pTG8.
ORGANISM	Cloning vector pTG8
REFERENCE	artificial sequences; vectors.
AUTHORS	1 (bases 1 to 3417)
JOURNAL	yang, Y. and Spector, A.
TITLE	Improved cloning vectors for transgene construction
MEDLINE	97313414
PUBMED	9187746
REFERENCE	2 (bases 1 to 3417)
AUTHORS	Yang, Y. and Spector, A.
TITLE	Direct Submission
JOURNAL	Submitted (21-MAY-1999)
FEATURES	Institute of Molecular Biology, University of Hong Kong, 8 Sassoon, Pokfulam, Hong Kong
source	Location/Qualifiers 1. ..3417 /organism="Cloning vector pTG8" /db_xref="taxon:119062"

```

/note="derived from Strategene vector pBS KS(+) by
insertion of rare FseI and PacI restriction sites flanking
the multiple cloning site"
complement(241..834)
/codon

```

```
complement(241, .834)
/codon_start=1
/transl_table=11
/product="lacZ"
/protein_id="AAF61634.1"
/db_xref="GI:7384996"
```

translation="WTMTPSAQGLTITKGNKSWVGPSPSRSTVSTLSINSQSPG
PLVLERPPWNSNSPYSSYILNARSLAVLQRDWNENPGVTQLNRLAHHPPFAS
PSQURLSINENWDAPCSAALSAAQGVVTVRSVTATLASALAPAPFAFF
SESLATFAGFPRLNLRGLPFLGRFSALRHLDPKKLD"
complement(2316..2867)

```
codon_start=1
transl_table=11
```

```
product="chloramphenicol acetyltransferase"  
protein_id="AAF61635.1"  
db_xref="GI:7384997"
```

Translation: "MEKKITGYTVDISOWHKEHFAQSQVACTNYOTVOLDITAF
TSTVKNKHKPYPAFHILARLNAHPERLMAKDGELVINDSWPCYTVFHEGTET
SILMSYEHDDPQZFLHYISODVACGENLAYEPKGFTEHNFVSNPWSFTSFDLNI
MNDNFAPSPWANIIRATRC"

NMDNEFAPSPWANIIRKATRC"
837 C 837 q 838 t

Query Match	35.2%	Score 352.6;	DB 12;	Length 3417;
Best Local Similarity	89.2%	Pred. No. 7.6e-92;		
Matches 355; Conservative				

[illegible]

Search completed: July 16, 2003, 13:27:48
Job time : 1825.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 03:41:41 ; Search time 170.989 seconds
(without alignments)
13183.634 Million cell updates/sec

Title: US-09-981-353-24

Perfect score: 1001

Sequence: 1 atgattacgaattcgagctc.....gtgagcaaaagccagcaaa 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185299 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	583	58.2	645	19	AAV63528
2	583	58.2	665	20	AAV63527
3	583	58.2	739	21	AAZ45659
4	579.8	57.9	718	19	AAV63527
5	545.8	54.5	560	18	AAZ45885
6	545.8	54.5	560	19	AAV16673
7	545.8	54.5	560	24	ABK15461
8	501	50.0	501	20	AAZ28676
9	429.4	42.9	661	22	AAK88295

C 10	355.8	35.5	5086	22	AAF84674	Nucleotide sequenc
C 11	352.4	35.2	5924	21	AAZ38921	hCAT1 binding huma
C 12	352.4	35.2	5925	21	AAZ38770	HCAT1 clone 25 ant
C 13	352.2	35.2	3789	21	AAZ39782	Plasmid pGB1MMN nu
C 14	352.2	35.2	5393	14	AAQ34611	Plasmid pr3. Synt
C 15	352	35.2	3983	22	AAQ10265	pUC/Ac vector DNA.
C 16	351.8	35.1	4824	17	AAZ08975	Shortened C1 gene
C 17	351.8	35.1	4824	20	AAZ15634	Nucleotide sequenc
C 18	351.8	35.1	6843	20	AAV83538	PCR-generated regi
C 19	351.8	35.1	7107	20	AAV83539	PCR-generated regi
C 20	351.8	35.1	8710	16	AAQ89650	pSC11 FIPV E1. Sy
C 21	351.8	35.1	9020	16	AAZ07652	pSC11 FIPV N. Syn
C 22	351.8	35.1	12011	20	AAV83540	PCR-generated regi
C 23	351.4	35.1	8062	22	AAZ04946	Plasmid pRK76 used
C 24	351.4	35.1	8153	22	AAZ04945	Plasmid pRK74 used
C 25	351.2	35.1	7958	24	AAI72490	H6-promoter human
C 26	351	35.1	8800	24	ABK95584	Yeast reporter con
C 27	351	35.1	10042	24	ABK95585	Yeast reporter con
C 28	350.8	35.0	584	22	AAF68704	Human lung tumour
C 29	350.8	35.0	584	24	ABK38615	cDNA encoding clon
C 30	350.8	35.0	655	24	ABK27689	Human colon cancer
C 31	350.8	35.0	738	20	AAZ09716	Plasmid pGEM DNA f
C 32	350.8	35.0	742	22	AAH82483	Human ovarian tumo
C 33	350.8	35.0	972	21	AAZ3524	T7 promoter sequen
C 34	350.8	35.0	2571	21	AAZ62824	Glyceroldehyde-3-p
C 35	350.8	35.0	2571	24	ABK17089	Eucalyptus grandis
C 36	350.8	35.0	3018	20	AAZ29905	Plasmid pGEM (RTM)
C 37	350.8	35.0	3216	21	AAZ49993	Plasmid pGNI for e
C 38	350.8	35.0	3216	22	AAZ06382	Vector pGNI DNA se
C 39	350.8	35.0	3253	19	AAV43439	DNA sequence of th
C 40	350.8	35.0	3323	24	AAZ27061	Plasmid pGNI DNA
C 41	350.8	35.0	3675	24	ABK12468	Plasmid vector pCX
C 42	350.8	35.0	3715	24	AAZ27064	Plasmid TopoRNAI D
C 43	350.8	35.0	3774	24	AAZ27062	Plasmid pGNS29 DNA
C 44	350.8	35.0	3968	22	AAZ09981	pHSP70-lmcs conatr
C 45	350.8	35.0	4001	24	AAZ27066	Plasmid pGNS9A DNA

ALIGNMENTS

RESULT 1
AAV63528
ID AAV63528 standard; cDNA; 645 BP.
XX
AC AAV63528;
XX
DT 28-JAN-1999 (first entry)
XX
DE Consensus sequence of CS110 sequence.
XX
KW EST sequence; CS110; gastrointestinal tract; cancer; adenocarcinoma;
KW Lymphoma; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 44..547
FT /*tag= a
XX
PN WO9844160-A1.
XX
PD 08-OCT-1998.
XX
PF 31-MAR-1998; 98WO-US06338.
XX
PR 31-MAR-1997; 97US-0829755.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
PI Gordon J, Granados EN, Hayden M, Hodges SC, Klass MR;
PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;

XX WPI: 1998-5571139/47.
 DR P-PSDB: AAW80477.
 XX New CS110 polypeptide(s) - useful for detecting gastrointestinal
 PT tract diseases, e.g. gastrointestinal tract cancers and to produce
 PT antibodies
 XX
 XX
 PS Claim 1; Page 91; 116pp; English.
 XX
 CC The present sequence is a consensus sequence of ESTs AAV63521-27
 CC which represent a set of contiguous and partially overlapping
 CC sequences designated CS110. The sequences are isolated from a cDNA
 CC library made from gastrointestinal tract tumour and normal tissues.
 CC The polypeptides are useful diagnostically to detect CS110
 CC antigen/anti-CS110 antibody in samples, e.g. to detect diseases
 CC and conditions of the gastrointestinal tract, especially
 CC cancers, e.g. adenocarcinoma and lymphoma.
 XX
 SQ Sequence 645 BP; 139 A; 174 C; 172 G; 160 T; 0 other;
 Query Match 58.2%; Score 583; DB 19; Length 645;
 Best Local Similarity 100.0%; Pred. No. 5.3e-130; Indels 0; Gaps 0;
 Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 CCCCAGAGTGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGC 86
 DB CCCCAGAGTGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGC 96
 QY 87 CATTGAGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGTGTGGAACGATT 146
 DB CATTGAGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGTGTGGAACGATT 156
 QY 147 CTCATCTCTGCAACCACTGGAGGCGCCCATCACCCTCCGCGTCCGAGTCAACAC 206
 DB CTCATCTCTGCAACCACTGGAGGCGCCCATCACCCTCCGCGTCCGAGTCAACAC 216
 QY 207 ATACTACATCGTAGTCTTTCAGTGGCGCTATGGCAAGTGTGGAGCGACTATGTGGTGG 266
 DB ATACTACATCGTAGTCTTTCAGTGGCGCTATGGCAAGTGTGGAGCGACTATGTGGTGG 276
 QY 267 TCGCAACGAGACCTGGAGGAGATCTTCTGCACCTGGGGAATCAGTATCCAGTTTC 326
 DB TCGCAACGAGACCTGGAGGAGATCTTCTGCACCTGGGGAATCAGTATCCAGTTTC 336
 QY 327 TGGGAAGTACAGTGGTACCTGAAGAGCTGGTATTTGTGACAGACAGGCGCGTATCT 386
 DB TGGGAAGTACAGTGGTACCTGAAGAGCTGGTATTTGTGACAGACAGGCGCGTATCT 396
 QY 387 GTCTTTTGGGAAGACAGTGGCAGCAAGTTTCAATGCCCTCCCTTGCACCCCAACACCGT 446
 DB GTCTTTTGGGAAGACAGTGGCAGCAAGTTTCAATGCCCTCCCTTGCACCCCAACACCGT 456
 QY 447 GCTCGGCTTCATCAGTGGCGGTCTGGTTCCTCATCGATGCCATTGGCCTGCACTGGGA 506
 DB GCTCGGCTTCATCAGTGGCGGTCTGGTTCCTCATCGATGCCATTGGCCTGCACTGGGA 516
 QY 507 TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCTCTCTCTCTTGGCAGGGGCACTGTGA 566
 DB TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCTCTCTCTCTTGGCAGGGGCACTGTGA 576
 QY 567 TGAGGAGTAAAGTCCCTTATCACTAACCCCACTCCAAATGG 609
 DB TGAGGAGTAAAGTCCCTTATCACTAACCCCACTCCAAATGG 619

RESULT 2

AAAX28677

ID AAX28677 standard; cDNA; 665 BP.

XX AAX28677;

AC AAX28677;

XX AAX28677;

DT 26-AUG-1999 (first entry)

XX Clone HP01738 encoding a human secretory signal protein (2).
 DE
 XX
 KW Human; secretory signal protein sequence; cell membrane;
 KW proliferation; differentiation; carcinostatic agent; antigen;
 KW antibody; probe; hybridisation; gene therapy; HP01738; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 58..562
 FT /*tag= a
 FT /product= "secretory signal protein"
 XX
 PN W09918204-A2.
 XX
 PD 15-APR-1999.
 XX
 PF 05-OCT-1998; 98WO-JP04476.
 XX
 PR 08-OCT-1997; 97JP-0276268.
 XX
 PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 XX
 PI Kato S, Kobayashi M, Sekine S, Yamaguchi T;
 XX
 DR WPI: 1999-264020/22.
 DR P-PSDB: AAY03235.
 XX
 PT Human proteins with secretory signal sequences and nucleotide
 PT sequences
 XX
 PS Claim 4; Page 73-74; 84pp; English.
 XX
 CC This is the nucleotide sequence of a clone encoding a human
 CC secretory signal protein sequence, used in the method of the
 CC invention. All of the proteins exist in the cell membrane, so are
 CC considered to be proteins controlling the proliferation and
 CC differentiation of the cells. They may be useful as carcinostatic
 CC agents or as antigens for preparing antibodies against the proteins.
 CC The cDNAs can be used as probes for gene diagnosis and gene sources
 CC for gene therapy, as well as for large-scale expression of the
 CC proteins.
 XX
 SQ Sequence 665 BP; 151 A; 176 C; 175 G; 163 T; 0 other;
 Query Match 58.2%; Score 583; DB 20; Length 665;
 Best Local Similarity 100.0%; Pred. No. 5.3e-130; Indels 0; Gaps 0;
 Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 CCCCAGAGTGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGC 86
 DB CCCCAGAGTGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGC 111
 QY 87 CATTGAGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGTGTGGAACGATT 146
 DB CATTGAGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGTGTGGAACGATT 171
 QY 147 CTCATCTCTGCAACCACTGGAGGCGCCCATCACCCTCCGCGTCCGAGTCAACAC 206
 DB CTCATCTCTGCAACCACTGGAGGCGCCCATCACCCTCCGCGTCCGAGTCAACAC 231
 QY 207 ATACTACATCGTAGTCTTTCAGTGGCGCTATGGCAAGTGTGGAGCGCACTATGTGGTGG 266
 DB ATACTACATCGTAGTCTTTCAGTGGCGCTATGGCAAGTGTGGAGCGCACTATGTGGTGG 291
 QY 267 TCGCAACGAGACCTGGAGGAGATCTTCTCACCCTCCGCGGGAATCAGTATCCAGTTTC 326
 DB TCGCAACGAGACCTGGAGGAGATCTTCTCACCCTCCGCGGGAATCAGTATCCAGTTTC 351
 QY 327 TGGGAAGTACAGTGGTACCTGAAGAGCTGGTATTTGTGACAGACAGGCGCGTATCT 386
 DB TGGGAAGTACAGTGGTACCTGAAGAGCTGGTATTTGTGACAGACAGGCGCGTATCT 386

Db 352 TGGGAAGTACAAAGTGTACCTGAAGAAGCTGGTATTTGTGACAGACAAGGGCCGCTATCT 411
 QY 387 GTCTTTTGGGAAGACAGTGGGCACAAAGTTTCAATGCCGTCCCTTTGCACCCCAACACCGT 446
 Db 412 GTCTTTTGGGAAGACAGTGGGCACAAAGTTTCAATGCCGTCCCTTTGCACCCCAACACCGT 471
 QY 447 GTCTCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCGCATTTGGCCTGCACCTGGGA 506
 Db 472 GCTCCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCGCATTTGGCCTGCACCTGGGA 531
 QY 507 TGTCTTACCCCACTAGTGCAGCAGATGCTGAGCCCTCTCTCTGTCAGTGGCAGGGGCACCTGTGA 566
 Db 532 TGTCTTACCCCACTAGTGCAGCAGATGCTGAGCCCTCTCTCTGTCAGTGGCAGGGGCACCTGTGA 591
 QY 567 TCAGGAGTAAGAATCCCTTATCACTAAACCCCAATCCCAATGG 609
 Db 592 TGAGGAGTAAGAATCCCTTATCACTAAACCCCAATCCCAATGG 634

RESULT 3

AA245659
 ID AA245659 standard; DNA; 739 BP.
 XX
 AC AA245659;
 XX

06-APR-2000 (first entry)

DNA sequence of colon specific gene (CSG) clone 1285234.

Human; colon specific gene; CSG; colon; metastasis; colon cancer;
 stomach cancer; small intestine cancer; ss.

Homo sapiens.

W09960161-AL.

25-NOV-1999.

12-MAY-1999; 99WO-0510498.

21-MAY-1998; 98US-0086266.

(DIAD-) DIADEXUS LLC.

Macina RA, Yang F, Sun Y;

WPI; 2000-126383/11.

Diagnosing, monitoring and staging colon cancer -

Example 1; Page 37; 47pp; English.

The present sequence represents a clone of a human colon specific gene (CSG) (gene ID number 239588). mRNA expression of the CSGs is highly specific for the colon. The specification describes a method for diagnosing the presence, or metastasis, of colon cancer in a patient. The method comprises measuring CSG levels in a cell, tissue or bodily fluid sample of the patient and a control (i.e. a human with no cancer), where increased CSG levels in the patient compared to the control is associated with the presence, or metastasis, of colon cancer. The methods may also be used for staging colon cancer, or for monitoring colon cancers, particularly of the colon, stomach and small intestine.

Sequence 739 BP; 167 A; 190 C; 183 G; 177 T; 22 other;

Query Match 58.2%; Score 583; DB 21; Length 739;
 Best Local Similarity 100.0%; Pred. No. 5.4e-130;
 Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CCCAGAAAGTTGACAGTGCCTCTCCTAGCCCTTCTGTGCGCTAGCCTCTGGCAATGC 86
 Db 80 CCCAGAAAGTTGACAGTGCCTCTCCTAGCCCTTCTGTGCGCTAGCCTCTGGCAATGC 139

QY 87 CATTTCAGGCCAGGCTTCTCTCTATAGTGGAGTATGGAAGTGGTGGTGGAAACCGATT 146
 Db 140 CATTTCAGGCCAGGCTTCTCTCTATAGTGGAGTATGGAAGTGGTGGTGGAAACCGATT 199
 QY 147 CTCCTATTCTGCAACACAGTTGGAGCGGCCCATCACCGCCCTCCGGGTCGAGTCAACAC 206
 Db 200 CTCCTATTCTGCAACACAGTTGGAGCGGCCCATCACCGCCCTCCGGGTCGAGTCAACAC 259
 QY 207 ATACTACATCGTGTAGTCTTCAGTGGCTGTATGGCAAGTGTGGAGCGACTATGTGGTGG 266
 Db 260 ATACTACATCGTGTAGTCTTCAGTGGCTGTATGGCAAGTGTGGAGCGACTATGTGGTGG 319
 QY 267 TCGCAACGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTGAATCCAGTTC 326
 Db 320 TCGCAACGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTGAATCCAGTTC 379
 QY 327 TGGGAAGTACAGTGGTACCTGAAGAAGCTGTATTTGTGACAGACAAGGGCCGCTATCT 386
 Db 380 TGGGAAGTACAGTGGTACCTGAAGAAGCTGTATTTGTGACAGACAAGGGCCGCTATCT 439
 QY 387 GTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGT 446
 Db 440 GTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCCGTCCCTTGCACCCCAACACCGT 499
 QY 447 GCTCCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCGCATTTGGCCTGCACCTGGGA 506
 Db 500 GCTCCGCTTCATCAGTGGCCGCTCTGGTTCTCTCATCGATGCGCATTTGGCCTGCACCTGGGA 559
 QY 507 TGTTCACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGCACCTGTGA 566
 Db 560 TGTTCACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGCACCTGTGA 619
 QY 567 TCAGGAGTAAGAATCCCTTATCACTAAACCCCAATCCCAATGG 609
 Db 620 TCAGGAGTAAGAATCCCTTATCACTAAACCCCAATCCCAATGG 662

RESULT 4

AAV63527
 ID AAV63527 standard; cDNA; 718 BP.
 XX
 AC AAV63527;
 XX

28-JAN-1999 (first entry)

Full length sequence of clone 609120IH corresponding to CS110 sequence.

EST sequence; CS110; gastrointestinal tract; cancer; adenocarcinoma;
 lymphoma; ss.

Homo sapiens.

W09844160-AL.

08-OCT-1998.

31-MAR-1998; 98WO-US06338.

31-MAR-1997; 97US-0829755.

(ABBO) ABBOTT LAB.

Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;
 Gordon J, Granados EN, Hayden M, Hodges SC, Klass MR;
 Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;
 WPI; 1998-557139/47.

New CS110 polypeptide(s) - useful for detecting gastrointestinal tract diseases, e.g. gastrointestinal tract cancers and to produce antibodies

PS Claim 1; Page 91; 116pp; English.
 CC AAV63521-27 represent a set of contiguous and partially overlapping EST
 CC sequences designated CS110. The sequences are isolated from a cDNA
 CC library made from gastrointestinal tract tumour and normal tissues.
 CC The polypeptides are useful diagnostically to detect CS110
 CC antigen/anti-CS110 antibody in samples, e.g. to detect diseases
 CC and conditions of the gastrointestinal tract, especially
 CC cancers, e.g. adenocarcinoma and lymphoma.
 XX
 XX Sequence 718 BP; 154 A; 202 C; 186 G; 176 T; 0 other;

Query Match 57.9%; Score 579.8; DB 19; Length 718;
 Best Local Similarity 99.7%; Pred. No. 3.1e-129;
 Matches 581; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 27 CCCAGAAATGTTGACAGTGGTCTCTAGCCCTTCTGTGCTCCAGCCCTCTGCAATGC 86
 DB 109 CCCAGAAATGTTGACAGTGGTCTCTAGCCCTTCTGTGCTCCAGCCCTCTGCAATGC 168
 QY 87 CATTGAGCCAGGCTCTCTCTATAGTGGAGATGATGGAAGTGGTGGTGAAGCCGATT 146
 DB 169 CATTGAGCCAGGCTCTCTCTATAGTGGAGATGATGGAAGTGGTGGTGAAGCCGATT 228
 QY 147 CTCATCTTCTGGCAACAGTTGGAGCCGCCATCACCAGCCCTCCGGGTCCGAGTCAACAC 206
 DB 229 CTCATCTTCTGGCAACAGTTGGAGCCGCCATCACCAGCCCTCCGGGTCCGAGTCAACAC 288
 QY 207 ATACTACATCTAGTCTTCAGTGGCGGTATGGCAAGTGGAGCGACTATGTGGTGG 266
 DB 289 ATACTACATCTAGTCTTCAGTGGCGGTATGGCAAGTGGAGCGACTATGTGGTGG 348
 QY 267 TCGCAACCGGAGACCTGGAGGAGATCTTCTGCACCCCTGGGAAATCAGTCCAGGTTTC 326
 DB 349 TCGCAACCGGAGACCTGGAGGAGATCTTCTGCACCCCTGGGAAATCAGTCCAGGTTTC 408
 QY 327 TGGGAAGTACAGTGGTACCTGAGAGACCTGGTATTTGTGACAGAGGGCCGCTATCT 386
 DB 409 TGGGAAGTACAGTGGTACCTGAGAGACCTGGTATTTGTGACAGAGGGCCGCTATCT 468
 QY 387 GTCCTTTGGGAAGACAGTGGCAGCAAGTTTCAATGCCCTCCCTTTGCACCCCAACACCGT 446
 DB 469 GTCCTTTGGGAAGACAGTGGCAGCAAGTTTCAATGCCCTCCCTTTGCACCCCAACACCGT 528
 QY 447 GCTCGCTTTCATAGTGGCGGTCTGTGTTCTCTCATCGATGCCATTTGGCCCTGCACCTGGGA 506
 DB 529 GCTCGCTTTCATAGTGGCGGTCTGTGTTCTCTCATCGATGCCATTTGGCCCTGCACCTGGGA 588
 QY 507 TGTTTACCCCACTAGCTGCAGAGATGCTGAGCCCTCTCTCTGCGGAGGGGCAGTGTGA 566
 DB 589 TGTTTACCCCACTAGCTGCAGAGATGCTGAGCCCTCTCTCTAGGAGGGGCAGTGTGA 648
 QY 567 TGAGGAGTGAAGACCTCCCTTATCACTAACCCGCCATCCAAATGG 609
 DB 649 TGAGGAGTGAAGACCTCCCTTATCACTAACCCGCCATCCATATGG 691

RESULT 5
 AAT45885
 ID AAT45885 standard; cDNA; 560 BP.
 XX
 AC AAT45885;
 XX
 DT 13-MAR-1997 (first entry)
 XX
 DE Human colon specific gene CS66 cDNA partial clone.
 XX
 KW Colon specific gene; CS66; colon cancer; metastasis; diagnosis;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT CDS 2...496
 FT /*tag= a
 PN WO9639419-A1.
 PD 12-DEC-1996.
 XX
 PF 06-JUN-1995; 95WO-US07289.
 XX
 PR 06-JUN-1995; 95WO-US07289.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Yu G;
 XX
 DR WPI: 1997-043054/04.
 DR P-PSDB; AAW06549.
 XX
 PT Human colon specific genes and their expression products - detection
 PT of which, in non-colon tissue samples, can be used as indication of
 PT colon cancer metastasis
 PS Claim 1; Fig 6; 60pp; English.
 XX
 CC 13 cDNA clones (AAU45880-92), 11 of them partial clones, correspond to
 CC human colon specific genes, designated CS61, CS62, etc., that are
 CC primarily expressed in tissues derived from the colon. CS67 and
 CC CS610 show reduced expression in colon cancer cells as compared to
 CC that in normal cells; the remaining genes are overexpressed in
 CC colon cancer. The partial cDNA sequences can be used to isolate
 CC full-length clones and genomic clones including the complete gene.
 CC CS6 nucleic acids can be used to produce CS6 polypeptides (see also
 CC AAW06545-53) in transformed host cells, as probes to detect disorders
 CC of the colon, partic. colon cancer and colon cancer metastasis, and
 CC in gene therapy.
 XX
 SQ Sequence 560 BP; 113 A; 157 C; 152 G; 138 T; 0 other;

Query Match 54.5%; Score 545.8; DB 18; Length 560;
 Best Local Similarity 99.5%; Pred. No. 4.1e-121;
 Matches 558; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 42 AGTCGCTCTCTAGCCCTTCTCTGCTCAGCTCTGCAATGCCATTCAGGCGCAGCTC 101
 DB 1 AGTCGCTCTCTAGCCCTTCTCTGCTCAGCTCTGCAATGCCATTCAGGCGCAGCTC 60
 QY 102 TTCCTCTCTATAGTGAGAGATGGAAGTGGTGGTGGAAAGCGATTCTCTCATCTCGCAA 161
 DB 61 TTCCTCTCTATAGTGAGAGATGGAAGTGGTGGTGGAAAGCGATTCTCTCATCTCGCAA 120
 QY 162 CCAGTTGGACGGCCCATCAGCCCTCCGGTCCGAGTCAACACATCTCATCTGAGG 221
 DB 121 CCAGTTGGACGGCCCATCAGCCCTCCGGTCCGAGTCAACACATCTCATCTGAGG 180
 QY 222 TCTTCAGTGGCTATGCAAGGTTGGAGGACTATGTGGTGGTCCGAGGAGACT 281
 DB 181 TCTTCAGTGGCTATGCAAGGTTGGAGGACTATGTGGTGGTCCGAGGAGACT 240
 QY 282 GGAGGAGATCTTCTGCACCTGGGAATCAGTATCAGGTTCTTGGGAAGTCAAGCTC 341
 DB 241 GGAGGAGATCTTCTGCACCTGGGAATCAGTATCAGGTTCTTGGGAAGTCAAGCTC 300
 QY 342 GTACCTGAAGAGCTGGTATTTGTGACAGAACGGCCGCTATCTGTCTTTTGGGAAGA 401
 DB 301 GTACCTGAAGAGCTGGTATTTGTGACAGAACGGCCGCTATCTGTCTTTTGGGAAGA 360
 QY 402 CAGTGGCAAGATTTCAATGGCTCCCTTGCACCCCAACACCGTCTCCGCTTCATCAG 461
 DB 361 CAGTGGCAAGATTTCAATGGCTCCCTTGCACCCCAACACCGTCTCCGCTTCATCAG 420
 QY 462 TGGCGGTCTGTCTCTCATCTGATGCCATTTGGCTGCACCTGGGATGTTTACCCCACTAG 521
 DB 421 TGGCGGTCTGTCTCTCATCTGATGCCATTTGGCTGCACCTGGGATGTTTACCCCACTAG 480

QY 522 CTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACCTGTGATGAGGAGTAAGAAGT 581
|||||
Db 481 CTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACCTGTGATGAGGAGTAAGAAGT 540
|||||
QY 582 CCTTATCACTAACCCCCATC 602
|||||
Db 541 -CCTTATCACTAACCCCCATC 560
|||||

RESULT 6

AAV16673
ID AAV16673 standard; cDNA; 560 BP.

XX AC AAV16673;
XX

DT 22-JUN-1998 (first entry)

DE Polynucleotide sequence of a colon-specific gene.

XX Colon-specific gene; probe: detection; expression; human;
KW diagnostic assay; colon cancer; antibody; screening; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers
FH 2..490
CDS /tag= a
FT /note= "no stop codon given"

FT

XX US5733748-A.

PN 31-MAR-1998.

XX

PD 06-JUN-1995; 95US-0469667.

XX

PF 06-JUN-1995; 95US-0469667.

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen C, Yu G;

XX WPI; 1998-229823/20.

XX P-PSDB; AAV46880.

DR Colon-specific nucleic acids - useful as probes for detecting colon

XX cancer micrometastases

PT Claim 15; Fig 6; 51pp; English.

XX

CC AAV16668-81 represent polynucleotide sequences of partial or full-length

CC cDNA clones of colon-specific genes. The polynucleotides can be used

CC as probes to detect expression of the corresponding human genes, e.g. in

CC diagnostic assays for detecting micrometastases of colon cancer.

CC Recombinant cells containing the polynucleotides can be used to

CC produce the polypeptides, in order that antibodies can be raised and

XX used in further screening or diagnostics.

XX Sequence 560 BP; 113 A; 157 C; 152 G; 138 T; 0 other;

QY Query Match 54.5%; Score 545.8; DB 19; Length 560;

Best Local Similarity 99.5%; Pred. No. 4.1e-121;

Matches 558; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 42 AGTCGCTCTCTAGCCCTCTCTGTGCTGCTCAGCCTCTGGCAATGCCATTGAGCCAGGTC 101
|||||
Db 1 AGTCGCTCTCTAGCCCTCTCTGTGCTGCTCAGCCTCTGGCAATGCCATTGAGCCAGGTC 60
|||||
QY 102 TTCTCTCTATAGTGGAGATGGAAGTGGTGGTGAAGCGATTCTCTCATTTCTGGCAA 161
|||||
Db 61 TTCTCTCTATAGTGGAGATGGAAGTGGTGGTGAAGCGATTCTCTCATTTCTGGCAA 120
|||||
QY 162 CCAGTTGGAGGGGCCCATCACCCTCCCGGGTCCGAGTCAACACATACATCTACATCTAGG 221
|||||

Db 121 CCAGTTGGAGGGGCCCATCACCCTCCCGGGTCCGAGTCAACACATACATCTAGG 180
|||||
QY 222 TCTTCAGGTGCGCTATGGCAAGGTGTGGAGCAGCTATGTGGTGTGCGCAACGGAGACCT 281
|||||
Db 181 TCTTCAGGTGCGCTATGGCAAGGTGTGGAGCAGCTATGTGGTGTGCGCAACGGAGACCT 240
|||||
QY 282 GGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTATCCAGGTTTCTGGGAAGTACAAGTG 341
|||||
Db 241 GGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTATCCAGGTTTCTGGGAAGTACAAGTG 300
|||||
QY 342 GTACCTGAAGAAGCTGGTATTTGTGACAGACAAGGCCCGCTATCTGTCTTTTGGGAAGA 401
|||||
Db 301 GTACCTGAAGAAGCTGGTATTTGTGACAGACAAGGCCCGCTATCTGTCTTTTGGGAAGA 360
|||||
QY 402 CAGTGGCACAAGTTTCAATGCGCTGCCCTTGACCCCAACACACCGTCTCCGCTTCATCAG 461
|||||
Db 361 CAGTGGCACAAGTTTCAATGCGCTGCCCTTGACCCCAACACACCGTCTCCGCTTCATCAG 420
|||||
QY 462 TGGCCGGTCTGTCTCTCATGATGCCATTGGCCCTGCACCTGGGATGTTTACCCCACTAG 521
|||||
Db 421 TGGCCGGTCTGTCTCTCATGATGCCATTGGCCCTGCACCTGGGATGTTTACCCCACTAG 480
|||||
QY 522 CTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACCTGTGATGAGGAGTAAGAAGT 581
|||||
Db 481 CTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACCTGTGATGAGGAGTAAGAAGT 540
|||||
QY 582 CCTTATCACTAACCCCCATC 602
|||||
Db 541 -CCTTATCACTAACCCCCATC 560
|||||

RESULT 7

ABK15461
ID ABK15461 standard; cDNA; 560 BP.

XX AC ABK15461;
XX

DT 08-MAY-2002 (first entry)

XX Human colon specific gene 6 (CSG6) cDNA.

DE Human; colon specific gene; CSG; cytostatic; colon disorder; gene;
KW gastric cancer; ulcerative colitis; granulomatous colitis; colon cancer;
KW cancer cell; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers
FH 2..490
CDS /tag= a
FT /product= "CSG6 protein"

FT /partial
FT /note= "No stop codon shown"

XX US6337195-B1.

PN 08-JAN-2002.

XX 31-MAR-1998; 98US-0224110.

XX 06-JUN-1995; 95US-0469667.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Yu G, Rosen C;

XX WPI; 2002-163239/21.

XX P-PSDB; AAU76209.

DR Human colon specific gene polypeptide, useful as diagnostic marker,

XX PT vaccine and for screening agonists and antagonists for treating colon

PT cancer

XX PS Example 1; Fig 6; 49pp; English.

XX CC This invention relates to the cDNA and protein sequences of novel

CC human colon specific genes. These genes which were isolated from a human

CC colon tissue cDNA library may have cytostatic activity and may be used

CC as a vaccine against colon disorders such as gastric cancer, ulcerative

CC colitis and granulomatous colitis. The nucleotide sequences of the

CC invention may be used to identify corresponding full length gene

CC sequences and is useful as diagnostic marker for colon cancer and as a

CC colon cancer vaccine. The protein sequences of the invention are

CC useful for preparing monoclonal antibodies specific to CSG proteins,

CC these antibodies can be used to isolate the protein from colon tissue or

CC for protein quantitation from a sample using enzyme linked immunosorbent

CC assay (ELISA). It is also useful for targeting cancer cells and for

CC screening agonist and antagonist for the CSG proteins which are useful

CC for treating colon cancer. The present sequence represents the human

CC colon specific gene sequence (CSG6) cDNA sequence of the invention.

XX SQ Sequence 560 BP; 113 A; 157 C; 132 G; 138 T; 0 other;

Query Match 54.5%; Score 545.8; DB 24; Length 560;

Best Local Similarity 99.5%; Pred. No. 4.1e-121;

Matches 558; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 42 AGTCGCTCTCTAGCCCTTCTCTGCTCAGCCCTCTGGCAATGCCATTCAGGCCAGGTC 101

DB 1 AGTCGCTCTCTAGCCCTTCTCTGCTCAGCCCTCTGGCAATGCCATTCAGGCCAGGTC 60

QY 102 TTCCTCTATAGTGAGAGTATGGAAGTGGTGGTGGAAAGCGATTCCTCATTCGGCAA 161

DB 61 TTCCTCTATAGTGAGAGTATGGAAGTGGTGGTGGAAAGCGATTCCTCATTCGGCAA 120

QY 162 CCAGTTGGACGGCCCATCATCCGCCCTCCCGGTCGAGTCAACACATCACTACCTGAGG 221

DB 121 CCAGTTGGACGGCCCATCATCCGCCCTCCCGGTCGAGTCAACACATCACTACCTGAGG 180

QY 222 TCTTCAGGTGGCTATGCAAGGTGTGGAGCGACTATGTGGGTGGTGGCAACGGAGACCT 281

DB 181 TCTTCAGGTGGCTATGCAAGGTGTGGAGCGACTATGTGGGTGGTGGCAACGGAGACCT 240

QY 282 GGAGGAGATCTTTCGACCCCTGGGATCATGATCAGTTCAGGTTCTGGGAAGTACAAGT 341

DB 241 GGAGGAGATCTTTCGACCCCTGGGATCATGATCAGGTTCTGGGAAGTACAAGT 300

QY 342 GTACCTGAAGAGCTGTGTTGTGACAGCAAGGCCGCTATCTGCTTTTGGGAAAGA 401

DB 301 GTACCTGAAGAGCTGTGTTGTGACAGCAAGGCCGCTATCTGCTTTTGGGAAAGA 360

QY 402 CAGTGGCACAAAGTTTCAATGCGCTCCCTTTCACCCCAACACCGTGTCCGCTTCATCAG 461

DB 361 CAGTGGCACAAAGTTTCAATGCGCTCCCTTTCACCCCAACACCGTGTCCGCTTCATCAG 420

QY 462 TGGCGGCTGTGGTCTCTCATCGATGCCATTTGGCGCTGCACATGGATGTTTACCCCACTAG 521

DB 421 TGGCGGCTGTGGTCTCTCATCGATGCCATTTGGCGCTGCACATGGATGTTTACCCCACTAG 480

QY 522 CTGACAGAGATGCTGAGCGCTCTCTCTTGGCAGGGGACATGTGATGAGGAGTAAGAAT 581

DB 481 CTGACAGAGATGCTGAGCGCTCTCTCTTGGCAGGGGACATGTGATGAGGAGTAAGAAT 540

QY 582 CCCTTATCACTAACCCCATC 602

DB 541 -CCTTATCACTAACCCCATC 560

RESULT 8

AAAX28676

ID AAX28676 standard; cDNA; 501 BP.

XX AAX28676;

AC AAX28676;

XX AAX28676;

DT 26-AUG-1999 (first entry)

XX DE Clone HP01738 encoding a human secretory signal protein (1).

XX KW Human; secretory signal protein sequence; cell membrane;

XX KW proliferation; differentiation; carcinostatic agent; antigen;

XX OS antibody; probe; hybridisation; gene therapy; HP01738; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..501

FT /*tag= a

FT /product= "secretory signal protein"

FT /note= "no termination codon given"

PN W09918204-A2.

XX 15-APR-1999.

XX 05-OCT-1998; 98WO-JP04476.

XX 08-OCT-1997; 97JP-0276268.

XX (PROT-) PROTEGENE INC.

XX (SAGA) SAGAMI CHEM RES CENT.

XX Kato S, Kobayashi M, Sekine S, Yamaguchi T;

XX WPI; 1999-264020/22.

XX P-PSDB; AAY03234.

XX Human proteins with secretory signal sequences and nucleotide

XX sequences

XX Claim 3; Page 71; 84pp; English.

XX This is the nucleotide sequence of a clone encoding a human

XX secretory signal protein sequence, used in the method of the

XX invention. All of the proteins exist in the cell membrane, so are

XX considered to be proteins controlling the proliferation and

XX differentiation of the cells. They may be useful as carcinostatic

XX agents or as antigens for preparing antibodies against the proteins.

XX The cDNAs can be used as probes for gene diagnosis and gene sources

XX for gene therapy, as well as for large-scale expression of the

XX proteins.

XX SQ Sequence 501 BP; 101 A; 136 C; 139 G; 125 T; 0 other;

Query Match 50.0%; Score 501; DB 20; Length 501;

Best Local Similarity 100.0%; Pred. No. 2.1e-110;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ATGTTGACAGTCGCTCTCTAGCCCTTCTCTGCTCAGCCTCTGGCAATGCCATTCAG 93

DB 1 ATGTTGACAGTCGCTCTCTAGCCCTTCTCTGCTCAGCCTCTGGCAATGCCATTCAG 60

QY 94 GCCAGGCTTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGAAAGCGATTCCTCAT 153

DB 61 GCCAGGCTTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGAAAGCGATTCCTCAT 120

QY 154 TCTGGCAACAGTTGGAGCGGCCCATCACGCCCTCCGGGTCCGAGTCAACACATCTAC 213

DB 121 TCTGGCAACAGTTGGAGCGGCCCATCACGCCCTCCGGGTCCGAGTCAACACATCTAC 180

QY 214 ATCTAGGTCTTTCAGGTGGCTATGGCAAGTGTGGAGCGACTATGTGGGTGGTGGCAAC 273

DB 181 ATCTAGGTCTTTCAGGTGGCTATGGCAAGTGTGGAGCGACTATGTGGGTGGTGGCAAC 240

QY 274 GGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGAATACAGTATCCAGGTTTCTGGGAAG 333

DB 241 GGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGAATACAGTATCCAGGTTTCTGGGAAG 300

QY 334 TACAAGTGGTACCTGAAGAGCTGGTATTTGTGACAGCAAGGCCGCTATCTCTTTT 393


```
Db 301 TACAAGTGTACTGAAGAAGCTGGTATTGTGCACACAAAGCGCGCTATCTGCTTTT 360
QY 394 GGAAGAAGACAGTGGCACAAGTTTCAATGCGGTCCCTTCACCCCAACACCGTGTCTCCGC 453
Db 361 GGAAGAAGACAGTGGCACAAGTTTCAATGCGGTCCCTTCACCCCAACACCGTGTCTCCGC 420
QY 454 TTTCATCAGTGGCGGTCTCTGCTTCTCTCATCGATGCCATGGCGTGCACCTGGATGTTTAC 513
Db 421 TTTCATCAGTGGCGGTCTCTGCTTCTCTCATCGATGCCATGGCGTGCACCTGGATGTTTAC 480
QY 514 CCCACTAGTGCAGCAGATGC 534
Db 481 CCCACTAGTGCAGCAGATGC 501

RESULT 9
AAK88295
ID AAK88295 standard; cDNA; 661 BP.
XX
AC
XX
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 611.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
```

PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-502630/55.
 DR P-PSDB; AAW92522.
 XX
 PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Claim 1; SEQ ID NO 611; 986pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
 CC system antigen of the invention.
 XX
 SQ Sequence 661 BP; 146 A; 175 C; 180 G; 156 T; 4 other;
 Query Match 42.9%; Score 429.4; DB 22; Length 661;
 Best Local Similarity 93.6%; Pred. No. 2.8e-93;
 Matches 468; Conservative 0; Mismatches 29; Indels 3; Gaps 2;
 QY 27 CCCAGAGATGTTGACAGTCGCTCTCCTAGCCCTTCTCTGTCCTCAGCTCTGGCAATGC 86
 DB 159 CCCAGAGATGTTGACAGTCGCTCTCCTAGCCCTTCTCTGTCCTCAGCTCTGGCAATGC 218
 QY 87 CATTACGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGGTGGGAAAGCGATT 146
 DB 219 CATTACGCCAGGCTTCTCTCTATAGTGGAGATGATGGAAGTGGTGGGAAAGCGATT 278
 QY 147 CTCATCTCTGGCAACAGTTGGAGCGCCCATCAGCCCTCCGGTCCGAGTCAACAC 206
 DB 279 CTCATCTCTGGCAACAGTTGGAGCGCCCATCAGCCCTCCGGTCCGAGTCAACAC 338
 QY 207 ATACTACATCTAGTCTTTCAGGTGCGCTATGGAAGTGGAGCGACTATGTGGTGG 266
 DB 339 ATACTACATCTAGTCTTTCAGGTGCGCTATGGAAGTGGAGCGACTATGTGGTGG 398
 QY 267 TCGCAACGGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTATCCAGGTTTC 326

Db 399 TCGCAACGGAGACCTGGAGGAGATCTTTCTGCACCCCTGGGGAATCAGTATCCAGGTTTC 458
 QY 327 TGGGAAGTACAAAGTGGTACCTGAAGAAGCTGGTATTTGTGACACAGAGGCCCGCTATCT 386
 Db 459 TGGGAAGTACAAAGTGGTACCTGAAGAAGCTGGTATTTGTGACAGAGAAGGCCCGCTATCT 518
 QY 387 GTCTTTTGGGAAAGACAGTGGCACAAGTTTCAATGCGCTGCCCTTGACACCCCAACACCGT 446
 Db 519 GTCTTTTGGGAAAGACAGTGGCACAAGTTTCAATGCGCTGCCCTTGACACCCCAACACCGT 578
 QY 447 GCTCCGCTTCATCAGTGGCGGCTCTGCTCTCTCA--TCGATGCCATTGGCTCGCAC--TG 503
 Db 579 CTTCGCTTCATCAATGCGCGGCTGGTCTCTTTCATGATGGCCATTGGNCTGCACCTG 638
 QY 504 GGATGTTTACCCCACTAGCT 523
 Db 639 GGATGTTTACCCCACTAGCT 658
 RESULT 10
 AAF84674/C
 ID AAF84674 standard; DNA: 5086 BP.
 AC AAF84674;
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE Nucleotide sequence of plasmid pMTet1.
 XX
 KW pMTet1; cytotoxic gene; selection; host cell isolation; ss.
 XX
 OS Synthetic.
 XX
 PN WO200125451-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 22-SEP-2000; 2000WO-GB03655.
 XX
 PR 01-OCT-1999; 99US-0157072.
 PR 25-JAN-2000; 2000GB-0001716.
 PR 07-SEP-2000; 2000US-0656503.
 XX
 PA (TRAN-) TRANSGENOMIC INC.
 XX
 PI Hornby DRJ, Martin MM;
 XX
 DR WPI; 2001-273580/28.
 XX
 PT One step cloning vector for quicker cloning of nucleic acid fragments
 PT comprises both a positive selection gene and a gene which protects
 PT against a selective agent -
 XX
 PS Claim 23; Page 41-42; 49pp; English.
 XX
 CC The present sequence represents plasmid pMTet1, which is used in the
 CC method of the invention. The specification describes one-step,
 CC non-plating methods for isolating transformed host cells which contain
 CC a vector incorporating a nucleic acid insert encoding a cytotoxic gene
 CC and a gene which provides the transformed host cell with a protective
 CC effect against the selection agent. The method uses a vector which
 CC comprises a cytotoxic gene permitting positive selection arranged so
 CC that it is insertionally inactivated by the incorporation into the vector
 CC of a heterologous nucleic acid insert; and a gene which provides a host
 CC cell transformed with the vector with a protective effect against the
 CC selection agent, but which does not significantly degrade the selection
 CC agent within a medium in which the host is present. The method allows
 CC the isolation of a transformed host cell which contains a vector
 CC incorporating a nucleic acid insert encoding a cytotoxic gene and a
 CC gene which provides the transformed host cell with a protective effect
 CC against the selection agent.
 XX

[illegible]

Db	2197	TTGTTATCCGCTCACAATTCCACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG	2138
Qy	721	GGGTGCCCTAATGAGTGAAGTAACTCACATTAAATTGCGTTGGGCTCACTGCCCCGCTTTTCCA	780
Db	2137	GGGTGCCCTAATGAGTGAAGTAACTCACATTAAATTGCGTTGGGCTCACTGCCCCGCTTTCCA	2078
Qy	781	GTCGGGAAACCTGTCTGTGCCAGCTGCATTAAATGAATCGGCCAACGCGCGGGGAGAGCGG	840
Db	2077	GTCGGGAAACCTGTCTGTGCCAGCTGCATTAAATGAATCGGCCAACGCGCGGGGAGAGCGG	2018
Qy	841	TTTCGGTATTGGGGGCTCTTCGGGTTCTCGCTCACTGCTCGTGGCGCTCGGTCGTTTCG	900
Db	2017	TTTCGGTATTGGGGGCTCTTCGGGTTCTCGCTCACTGCTCGTGGCGCTCGGTCGTTTCG	1958
Qy	901	GCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATTCCACAGAATCAGG	960
Db	1957	GCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATTCCACAGAATCAGG	1898
Qy	961	GGGATAGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA	1001
Db	1897	GGATAAGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA	1857
RESULT 12			
AAZ38770/c			
ID	AAZ38770 standard; DNA; 5925 bp.		
XX			
AC	AAZ38770;		
DT	08-FEB-2000 (first entry)		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 08:51:42 ; Search time 1092.78 Seconds

(without alignments)
14835.236 Million cell updates/sec

Title: US-09-981-353-24

Perfect score: 1001

Sequence: 1 atgattagcaattcgatctc.....gtgagcaaaagccagcaaa 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esti:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	575	57.4	605	9 AI732905	AI732905 of72a12.x
c 2	544.4	54.4	581	13 BM021151	BM021151 ie76h03.x
c 3	542.8	54.2	581	13 BM023112	BM023112 ie79e10.x
c 4	542.4	54.2	602	9 AI833168	AI833168 at75h09.x
c 5	530	52.9	549	14 BM934914	BM934914 UI-M-BH3-
c 6	521.4	52.1	633	13 BI763402	BI763402 603047590

7	517.8	51.7	572	13	BM021411	BM021411 ie76h03.y
c 8	497.8	46.9	566	9	AI991154	AI991154 wu38h12.x
c 9	469.4	46.9	520	14	BQ028047	BQ028047 UI-H-C00-
c 10	447	44.7	490	14	BM987195	BM987195 UI-H-C00-
c 11	441	44.1	485	14	BM987408	BM987408 UI-H-C00-
c 12	435.4	43.5	492	14	BM987303	BM987303 UI-H-C00-
c 13	435.2	43.5	497	9	AA828206	AA828206 of05d11.s
c 14	426.4	42.6	455	9	AI749030	AI749030 at38b05.x
c 15	423.2	42.3	467	10	AW001526	AW001526 wu34a06.x
c 16	419.6	41.9	489	13	BM023370	BM023370 ie79e10.y
c 17	418.4	41.8	448	9	AI984533	AI984533 wu22h06.x
c 18	416.8	41.6	449	9	AI880265	AI880265 at47007.x
c 19	410.4	41.0	456	14	BM987839	BM987839 UI-H-C00-
c 20	409.8	40.9	496	10	AW000842	AW000842 wu46c02.x
c 21	400.4	40.0	466	10	AW001333	AW001333 wu28h11.x
c 22	394.4	39.4	464	9	AA937899	AA937899 of72a12.s
c 23	393	39.3	446	9	AI984522	AI984522 wu22g02.x
c 24	386.2	38.6	440	10	AW456570	AW456570 UI-M-BH3-
c 25	382.4	38.2	427	10	AW001511	AW001511 wu33g01.x
c 26	375.2	37.5	467	10	AW844215	AW844215 RC4-CN004
c 27	374	37.4	418	14	BM987670	BM987670 UI-H-C00-
c 28	370.4	37.0	683	12	BF585282	BF585282 602101945
c 29	370.4	37.0	743	14	BQ960080	BQ960080 AGENCOURT
c 30	370.2	37.0	550	12	BG879223	BG879223 lb60d09.y
c 31	370.2	37.0	564	12	BG311182	BG311182 lb21a09.y
c 32	370.2	37.0	647	13	BQ967068	BQ967068 602833913
c 33	370.2	37.0	662	14	BQ952658	BQ952658 AGENCOURT
c 34	370.2	37.0	666	12	BF583056	BF583056 602097944
c 35	370.2	37.0	666	13	BQ967428	BQ967428 602833573
c 36	370.2	37.0	666	14	BQ954539	BQ954539 AGENCOURT
c 37	370.2	37.0	668	14	BQ921513	BQ921513 AGENCOURT
c 38	370.2	37.0	674	12	BF577589	BF577589 602092454
c 39	370.2	37.0	675	14	BQ961186	BQ961186 AGENCOURT
c 40	370.2	37.0	676	14	BQ947379	BQ947379 AGENCOURT
c 41	370.2	37.0	677	14	BQ929155	BQ929155 AGENCOURT
c 42	370.2	37.0	679	12	BF579244	BF579244 602093414
c 43	370.2	37.0	686	13	BQ966346	BQ966346 602832832
c 44	370.2	37.0	689	14	BQ939500	BQ939500 AGENCOURT
c 45	370.2	37.0	689	14	BQ945842	BQ945842 AGENCOURT

ALIGNMENTS

RESULT 1
AI732905/c
LOCUS AI732905 605 bp mRNA linear EST 13-DEC-1999
DEFINITION of72a12.x5 NCI_CGAP_Co8 Homo sapiens CDNA clone IMAGE:1433870 3' similar to TR:O60844 O60844 HOMOLOG OF RAT ZYMOGEN GRANULE MEMBRANE PROTEIN. ; mRNA sequence.
ACCESSION AI732905 GI:5054018
VERSION AI732905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 605)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Prepared by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: National Cancer Institute, Cancer Genome
Anatomy Project (CGAP), Tumor Gene Index
This read has been verified (found to hit its original self in the
correct orientation)

Insert Length: 725 Std Error: 0.00

Seq primer: -400P from Gluco

High quality sequence stop: 469.

FEATURES

source
1. 605
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1435870"
/clone_lib="NCI-CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified p7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
148 a 160 c 160 g 136 t 1 others

BASE COUNT
ORIGIN

Query Match 57.4%; Score 575; DB 9; Length 605;

Best Local Similarity 99.8%; Pred. No. 1.5e-154;

Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 34 ATGTTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGCCATTACG 93

Db 605 ATGTTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGCCATTACG 546

QY 94 GCCAGTCTTCCCTCTAGTGGAGATGATGGAAGTGGTGGAAACCGATTCTCTCAT 153

Db 545 GCCAGTCTTCCCTCTAGTGGAGATGATGGAAGTGGTGGAAACCGATTCTCTCAT 486

QY 154 TCTGGCAACCACTGTTGGAGCGGCCCATCCGCCCTCCGGGTCCGAGTCAACACATATAC 213

Db 485 TCTGGCAACCACTGTTGGAGCGGCCCATCCGCCCTCCGGGTCCGAGTCAACACATATAC 426

QY 214 ATCGTAGTCTTCAGTGCGCTATGGCAAGTGTGGAGCGACTATGTGGGTGGCAAC 273

Db 425 ATCGTAGTCTTCAGTGCGCTATGGCAAGTGTGGAGCGACTATGTGGGTGGCAAC 366

QY 274 GGAGACCTGGAGGAGATCTTCTGCACCCTGGGAATCAGTGATCCAGGTTTCTGGGAAG 333

Db 365 GGAGACCTGGAGGAGATCTTCTGCACCCTGGGAATCAGTGATCCAGGTTTCTGGGAAG 306

QY 334 TACAAGTGTACCTGAAGAAGCTGGTATTGTGACAGACAGGCGCGCTATCTGCTTTT 393

Db 305 TACAAGTGTACCTGAAGAAGCTGGTATTGTGACAGACAGGCGCGCTATCTGCTTTT 246

QY 394 GGGAAAGACAGTGCACAAAGTTTCAATGCCCTCCCTTTCACCCCAACACCGTCTCCGC 453

Db 245 GGGAAAGACAGTGCACAAAGTTTCAATGCCCTCCCTTTCACCCCAACACCGTCTCCGC 186

QY 454 TTCATCAGTGGCCGGTCTGTTCTCTCATGCCATGCCATGGCCCTGCACCTGGGATGTTAC 513

Db 185 TTCATCAGTGGCCGGTCTGTTCTCTCATGCCATGCCATGGCCCTGCACCTGGGATGTTAC 126

QY 514 CCCACTAGCTGCACGATGCTGAGCCCTCCCTCTTGGCAGGGGCACTGTGATGAGGAG 573

Db 125 CCCACTAGCTGCACGATGCTGAGCCCTCCCTCTTGGCAGGGGCACTGTGATGAGGAG 66

QY 574 TAAGAAGTCCCTTATCACTAACCCCAATCCAAATGG 609

Db 65 TAAGAAGTCCCTTATCACTAACCCCAATCCAAATGG 30

RESULT 2

BM021151/c
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 581

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5673053"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;

Site 2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size-selected by column

fractionation; average insert size 1.08 kb. Library was

amplified once on solid support and plasmid DNA from

library was prepared. The library DNA was normalized by

method #4 from Bonaldo, Lennon, and Soares 1996 Genome

Research 6:791-806; 0.5 microgram single-stranded library

plasmid DNA was mixed with 5 micrograms PCR product

representing library inserts and hybridized to an Ecot of

20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this

library."

BASE COUNT 139 a 155 c 150 g 137 t

ORIGIN

Query Match 54.4%; Score 544.4; DB 13; Length 581;

Best Local Similarity 99.8%; Pred. No. 9.9e-146;

Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 64 TGTGCCTCAGCCTCTGGCAATGCCATTTCAGCCAGGTCCTTCTCTATAGTGGAGATAT 123

Db 581 TGTGCCTCAGCCTCTGGCAATGCCATTTCAGCCAGGTCCTTCTCTATAGTGGAGATAT 522

QY 124 GGAAGTGTGTGGAAAGCGATTCTCTCATTTCTGGCAACCAAGTTGGACGCCCATCACC 183

Db 521 GGAAGTGTGTGGAAAGCGATTCTCTCATTTCTGGCAACCAAGTTGGACGCCCATCACC 462

184 GCCCTCCGGTCCGAGTCAACATACATCTAGTCTTTCAGTGGCTATGGCAAG 243
|||||
461 GCCCTCCGGTCCGAGTCAACATACATCTAGTCTTTCAGTGGCTATGGCAAG 402
|||||
244 GTGTGGAGCGACTATGTGGTGGTGGCGAGACCTGGAGGAGATCTTTCTGCACCT 303
Db |||||||
401 GTGTGGAGCGACTATGTGGTGGTGGCGAGACCTGGAGGAGATCTTTCTGCACCT 342
|||||
304 GGGGAATCAGTATCCAGTTTCTGGGAAGTACAAGTGTACTCTGAAGAGCTGTATTT 363
Db |||||||
341 GGGGAATCAGTATCCAGTTTCTGGGAAGTACAAGTGTACTCTGAAGAGCTGTATTT 282
|||||
364 GTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCC 423
Db |||||||
281 GTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCC 222
|||||
424 GTCCCTTGACCCCAACACCGTGTCCGCTTCATCAGTGGCGGTCTGTCTCTCATC 483
Db |||||||
221 GTCCCTTGACCCCAACACCGTGTCCGCTTCATCAGTGGCGGTCTGTCTCTCATC 162
|||||
484 GATGCATTGGCTGCACCTGAGTCTTACCCCTAGCTGCAGCAGATGCTGAGCCTCC 543
Db |||||||
161 GATGCATTGGCTGCACCTGAGTCTTACCCCTAGCTGCAGCAGATGCTGAGCCTCC 102
|||||
544 TCTCTTGAGCGGCGACTGTGATGAGGAGTAAAGTCCCTTATCATAACCCCATCC 603
Db |||||||
101 TCTCTTGAGCGGCGACTGTGATGAGGAGTAAAGTCCCTTATCATAACCCCATCC 42
|||||
604 AAATGG 609
Db |||||||
41 AAATGG 36

RESULT 3

BM023112/c

LOCUS

DEFINITION
581 bp mRNA linear EST 12-MAR-2002
cdna clone IMAGE:5673306 3' similar to TR:060844 060844 HOMOLOG OF
RAT ZYMOGEN GRANULE MEMBRANE PROTEIN. ; mRNA sequence.

ACCESSION

BM023112

VERSION

BM023112.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota

Metazoa

Chordata

Craniata

Vertebrata

Euteleostomi

Mammalia

Eutheria

Primates

Catarrhini

Hominoidea

Homo

1 (bases 1 to 581)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,

Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas

M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,

Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brownjefas.harvard.edu) This sequence now available from the IMAGE

consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 344.

FEATURES

source

I. .581

/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:5673306"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1;
Site:2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
method 44 from Bonaldo, Lennan, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT 140 a 155 c 149 g 137 t
ORIGIN

Query Match 54.2%; Score 542.8; DB 13; Length 581;

Best Local Similarity 99.6%; Pred. No. 2.8e-145;

Matches 544; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 64 TGTGCTCAGCTCTGGCAATGCCATTTCAGGCCAGGTCTTCTCTCTATAGTGGAGATAT 123

Db |||||||

581 TGTGCTCAGCTCTGGCAATGCCATTTCAGGCCAGGTCTTCTCTCTATAGTGGAGATAT 522

QY 124 GGAAGTGTGTGGAAAGCGATTCTCTCATCTTGCAACACAGTGGACGGCCCCATCACC 183

Db |||||||

521 GGAAGTGTGTGGAAAGCGATTCTCTCATCTTGCAACACAGTGGACGGCCCCATCACC 462

QY 184 GCCCTCCGGTCCGAGTCAACATACATCTAGTCTTTCAGTGGCTATGGCAAG 243

Db |||||||

461 GCCCTCCGGTCCGAGTCAACATACATCTAGTCTTTCAGTGGCTATGGCAAG 402

QY 244 GTGTGAGCGACTATGTGGTGGTGGCGAGACCTGGAGGAGATCTTTCTGCACCT 303

Db |||||||

401 GTGTGAGCGACTATGTGGTGGTGGCGAGACCTGGAGGAGATCTTTCTGCACCT 342

QY 304 GGGGAATCAGTATCCAGTTTCTGGGAAGTACAAGTGTACTCTGAAGAGCTGTATTT 363

Db |||||||

341 GGGGAATCAGTATCCAGTTTCTGGGAAGTACAAGTGTACTCTGAAGAGCTGTATTT 282

QY 364 GTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCC 423

Db |||||||

281 GTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAGTTTCAATGCC 222

QY 424 GTCCCTTGACCCCAACACCGTGTCCGCTTCATCAGTGGCGGTCTGTCTCTCATC 483

Db |||||||

221 GTCCCTTGACCCCAACACCGTGTCCGCTTCATCAGTGGCGGTCTGTCTCTCATC 162

QY 484 GATGCATTGGCTGCACCTGAGTCTTACCCCTAGCTGCAGCAGATGCTGAGCCTCC 543

Db |||||||

161 GATGCATTGGCTGCACCTGAGTCTTACCCCTAGCTGCAGCAGATGCTGAGCCTCC 102

QY 544 TCTCTTGAGCGGCGACTGTGATGAGGAGTAAAGTCCCTTATCATAACCCCATCC 603

Db |||||||

101 TCTCTTGAGCGGCGACTGTGATGAGGAGTAAAGTCCCTTATCATAACCCCATCC 42

QY 604 AAATGG 609

Db |||||||

41 AAATGG 36

RESULT 4

AI833168/c

LOCUS

DEFINITION

at75h09.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone

EST 13-JUL-1999

IMAGE:2377889 3' similar to TR:060844 060844 HOMOLOG OF RAT ZYMOGEN GRANULE MEMBRANE PROTEIN. ;, mRNA sequence.

ACCESSION AI833168

VERSION AI833168.1 GI:5455148

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 602)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maier, M., Martin

, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,

White, Y., Wylie, T., Waterston, R. and Wilson, R.

WashU-NCI human EST Project

Unpublished (1997)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (infoimage.llnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco

High quality sequence stop: 214.

Location/Qualifiers

1..602

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2377889"

/clone.lib="Barstead colon HPLRB7"

/sex="male"

/dev_stage="adult, age 25"

/lab_host="DH10B (phage resistant)"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site: EcoRI; Site 2: NotI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[5' AATTCATACTAAT 3' and 5' ATTACTAGT 3'], digested

with Not I and cloned into the Not I and Eco RI sites of

the modified pT7T3 vector. Library constructed by Bob

Barstead."

BASE COUNT 152 a 159 c 161 g 129 t 1 Others

ORIGIN

Query Match 54.2%; Score 542.4; DB 9; Length 602;

Best Local Similarity 98.4%; Pred. No. 3.8e-145;

Matches 568; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

34 ATGTTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGCCATTCAG 93

|||||

602 ATGTTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCCTCTGGCAATGCCATTCAG 544

|||||

94 GCCAGGTCTTCTCTATAGTGGAGAGTATGGAAGTGTGTGGAAACCGATTCTCTCAT 153

|||||

543 GCCAGGTCTTCTCTATAGTGGAGAGTATGGAAGTGTGTGGAAACCGATTCTCTCAT 484

|||||

154 TCTGGCAACCAAGTGTGGAGCGGCCCATCCGCCCTCCCGGGTCCGAGTCAACATATAC 213

|||||

483 TCTGGCAACCAAGTGTGGAGCGGCCCATCCGCCCTCCCGGGTCCGAGTCAACATATAC 424

|||||

214 ATCTAGTCTTTCAGTTCGCTATGCGAAGTGTGGAGCGGCTATGTGGGTGGTGGCAAC 273

|||||

423 ATCTAGTCTTTCAGTTCGCTATGCGAAGTGTGGAGCGGCTATGTGGGTGGTGGCAAC 364

|||||

274 GGAGACCTGGAGGAGATCTTTCTGCACCCTCGGGAATCAGTATCCAGGTTTCTGGGAAG 333

|||||

363 GGAGACCTGGAGGAGATCTTTCTGCACCCTCGGGAATCAGTATCCAGGTTTCTGGGAAG 304

|||||

334 TACAAGTGGTACCTGAAGAAGCTGGTATTTCTGACAGACAAGGG-CCGCTATCTGTCTTT 392

|||||

303 TACAAGTGGTACCTGAAGAAGCTGGTATTTGTGACAGACAAGGCCCTTTTGTCTTTT 244

|||||

393 TGGCAAAAGACAGTGGCACAAGCTTTCAATGCCGTCCTTGCACCCCAACACCGTGCCTCG 452

|||||

243 TGGAAAGACAGTGGCACAAGCTTTCAATGCCGTCCTTGCACCCCAACACCGTGCCTCG 184

|||||

453 CTTTCATCAGTGGCGGCTCTGTTCTCTCATCGATGCCATTCGCTCGCACTGGGATGTTTA 512

|||||

183 CTTTCATCAGTGGCGGCTTGGTTCTCTCATCGATGCCATTCGCTCGCACTGGGATGTTTA 124

|||||

513 CCCCACTAGTGGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGA 572

|||||

123 CCCCACTAGTGGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGGGCACTGTGATGAGGA 64

|||||

573 GTAAGAAGCTCCCTTATCACTAACCCCATCCCAATGG 609

|||||

63 GTAAGAAGCTCCCTTATCACTAACCCCATCCCAATGG 27

|||||

RESULT 5

BM934914

LOCUS 549 bp mRNA linear EST 13-MAR-2002

DEFINITION UI-M-BH3-aqs-e-10-0-UI.r1 NIH_BMAP_M.S4 Mus musculus cDNA clone

UI-M-BH3-aqs-e-10-0-UI 5', mRNA sequence.

BM934914

ACCESSION BM934914.1 GI:19394066

VERSION EST.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 549)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mst@mail.nih.gov

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 REVERSE.

Location/Qualifiers

1..549

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH3-aqs-e-10-0-UI"

/clone.lib="NIH_BMAP_M.S4"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site: Not I; Site 2: Eco RI; The

NIH_BMAP_M.S4 library is a subtracted library of a series,

ultimately derived from a mixture of individually tagged

normalized libraries from ten regions of the mouse brain

(cerebellum, brain stems, olfactory bulbs, hypothalamus,

cortex, amygdala, basal ganglia, pineal gland, striatum,

hippocampus) after a series of subtractions to reduce the

representation of cDNAs from which ESTs had already been

generated. The following serially subtracted libraries

were generated in this process: NIH_BMAP_M.S4,

NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,

NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library

(NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996).

```

BASE COUNT      124 a  144 c  149 g  132 t
ORIGIN

Query Match      52.9%; Score 530; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.4e-141;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GCAATGCCATTTCAGGCGAGGCTCTCTCTCTATAGTGAGAGTATGGAAGTGTGGGAA 139
    |||||||
Db 1 GCAATGCCATTTCAGGCGAGGCTCTCTCTCTATAGTGAGAGTATGGAAGTGTGGGAA 60
    |||||||
QY 140 AGCGATTCTCTATCTTGGCAACAGTGTGACGGCCCCATCACCGCCCTCCGGGTCCGAG 199
    |||||||
Db 61 AGCGATTCTCTATCTTGGCAACAGTGTGACGGCCCCATCACCGCCCTCCGGGTCCGAG 120
    |||||||
QY 200 TCAACACATACATACATGAGTCTTCAGTGGCGCTATGCGAAGTGTGGAGGACTATG 259
    |||||||
Db 121 TCAACACATACATACATGAGTCTTCAGTGGCGCTATGCGAAGTGTGGAGGACTATG 180
    |||||||
QY 260 TGGGTGGTGGCAACGAGACCTGGAGGAGATCTTCTGCACCTCTGGGGAATCAGTGATCC 319
    |||||||
Db 181 TGGGTGGTGGCAACGAGACCTGGAGGAGATCTTCTGCACCTCTGGGGAATCAGTGATCC 240
    |||||||
QY 320 AGGTTTCTGGAGTACAGTGTACTTGAAGAGTGTATTTGACAGACAAAGGGCC 379
    |||||||
Db 241 AGGTTTCTGGAGTACAGTGTACTTGAAGAGTGTATTTGACAGACAAAGGGCC 300
    |||||||
QY 380 GCTATCTGTCTTTTGGGAAGACAGTGGCACAGTTTCAATCCGCTCCCTTGCACCCCA 439
    |||||||
Db 301 GCTATCTGTCTTTTGGGAAGACAGTGGCACAGTTTCAATCCGCTCCCTTGCACCCCA 360
    |||||||
QY 440 ACACCTGCTCCGCTTCATCAGTGGCGGTCTGTCTCTCATCGATGCCATTTGCCCTGC 499
    |||||||
Db 361 ACACCTGCTCCGCTTCATCAGTGGCGGTCTGTCTCTCATCGATGCCATTTGCCCTGC 420
    |||||||
QY 500 ACTGGGATGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGC 559
    |||||||
Db 421 ACTGGGATGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGC 480
    |||||||
QY 560 ACTGTGATGAGGAGTAAGAACTCCCTTATCTACTAACCCCACTCCAAATGG 609
    |||||||
Db 481 ACTGTGATGAGGAGTAAGAACTCCCTTATCTACTAACCCCACTCCAAATGG 530
    |||||||

RESULT 6
BI763402
LOCUS
DEFINITION 603047590F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187790 5',
          mRNA sequence.
ACCESSION BI763402
VERSION BI763402.1 GI:15754980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

```

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11469 row: C column: 23
High quality sequence stop: 618.

FEATURES

Source

1.633
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187790"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."
BASE COUNT 150 a 169 c 167 g 146 t 1 Others
ORIGIN

```

Query Match      52.1%; Score 521.4; DB 13; Length 633;
Best Local Similarity 99.0%; Pred. No. 4.3e-139;
Matches 566; Conservative 2; Mismatches 2; Indels 4; Gaps 4;

QY 40 ACAGTCGCTCTCCCTAGCCCTT-CTCTGTGCTCAGCCCTCGCAATGCATTCAGGCGAG 98
    |||||||
Db 13 ACAGTCGCTCTCCCTAGCCCTTCTCTGTGCTCAGCCCTCGCAAGG-CATTTCAGGCGAG 71
    |||||||
QY 99 GTCTTCTCTCTATAGTGGAGAGTATGGAAGTGTGGTGGAAAGGATTCCTCTCATCTGG 158
    |||||||
Db 72 GTCTTCTCTCTATAGTGGAGAGTATGGAAGTGTGGTGGAAAGGATTCCTCTCATCTGG 131
    |||||||
QY 159 CAACAGTTGGACGGCCCCATCACGCCCTCCGGGTCCGAGTCAACATACATACATCCGT 218
    |||||||
Db 132 CAACAGTTGGACGGCCCCATCACGCCCTCCGGGTCCGAGTCAACATACATACATCCGT 191
    |||||||
QY 219 AGGCTTTCAGTGGCGCTATGGCAAGGTGTGGAGCGACTATGTGGGTGGTGGCAACGGAGA 278
    |||||||
Db 192 AGGCTTTCAGTGGCGCTATGGCAAGGTGTGGAGCGACTATGTGGGTGGTGGCAACGGAGA 251
    |||||||
QY 279 CCTGGAGAGATCTTCTGACCCCTGGGATCAGTATCCAGGTTTCTGGGAAGTACAA 338
    |||||||
Db 252 CCTGGAGAGATCTTCTGACCCCTGGGATCAGTATCCAGGTTTCTGGGAAGTACAA 311
    |||||||
QY 339 GTGGTACCTCAAGAAGCTGTGTTGTGACAGACAAGGCGCTATCTGCTTTTGGGAA 398
    |||||||
Db 312 GTGGTACCTCAAGAAGCTGTGTTGTGACAGACAAGGCGCTATCTGCTTTTGGGAA 371
    |||||||
QY 399 AGACAGTGGGCAACAAGTTTCAATCCGCTCCCTTGCACCCCAACACCGTGTCCGCTTCAT 458
    |||||||
Db 372 AGACAGTGGGCAACAAGTTTCAATCCGCTCCCTTGCACCCCAACACCGTGTCCGCTTCAT 431
    |||||||
QY 459 CAGTGGCGGTCTGGTTCTCTCATCGATGCCATGGCCTGCATGGGATGTTTACCCCA 518
    |||||||
Db 432 CAGTGGCGGTCTGGTTCTCTCATCGATGCCA-TGGCCTGCATGGGATGTTTACCCCA 490
    |||||||
QY 519 TAGCTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGCGACTGTGTATGAGGAGTAAGA 578
    |||||||
Db 491 TAGCTGCAGCAGATGCTGAGCCTCTCTCTTGGCAGGCGACTGTGTATGAGGAGTAAGA 550
    |||||||
QY 579 ACTCCC-TTATCTACTAACCCCACTCCAAATGG 609
    |||||||

```

```
||||| ACTCCCGTATCACTAACCCCAATG 582
Db 551

BM021411 572 bp mRNA linear EST 12-MAR-2002
ie76h03.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5673053 5' similar to TR:060844 060844 HOMOLOG OF
RAT ZYMOGEN GRANULE MEMBRANE PROTEIN. ; mRNA sequence.
BM021411
ACCESSION BM021411.1 GI:16535767
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 572)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvill,I.R., Williams,T.,
Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
COMMENT Other ESTs: ie76h03.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 430.
FEATURES
source
Location/Qualifiers
1..572
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5673053"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B",
/Note="Organ: Pancreas; Vector: pSPORT1; Site.1: Not 1;
Site.2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life technologies). cDNA
made by oligo-dr priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
BASE COUNT 118 a 154 c 155 g 145 t
ORIGIN
Query Match 51.7%; Score 517.8; DB 13; Length 572;
Best Local Similarity 99.6%; Pred. No. 4.5e-138;
Matches 519; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 27 CCCAGAAATGTTGACAGTCGCTCTCCCTAGCCCTCTCTGCGCTCAGCCCTCTGCGCAATGC 86
|||||
52 CCCAGAAATGTTGACAGTCGCTCTCCCTAGCCCTCTCTGCGCTCAGCCCTCTGCGCAATGC 111
|
87 CATTCAGGCCAGGCTCTCCCTCTATAGTGGAGATATGGAAGTGGTGGTGAAGCCGATT 146
|
112 CATTCAGGCCAGGCTCTCCCTCTATAGTGGAGATATGGAAGTGGTGGTGAAGCCGATT 171
|
147 CTCATCTTCGGCAACCAAGTTGGACGCCCCATCACGCCCTCCGGTCCGAGTCAACAC 206
|
172 CTCATCTTCGGCAACCAAGTTGGACGGTCTCATCCGCCCTCCGGTCCGAGTCAACAC 231
|
207 ATACTACATCGTAGGCTCTTCAGGTGCGCTATGGCAAGGTGTGGAGCACTATGTGGGTGG 266
|
232 ATACTACATCGTAGGCTCTTCAGGTGCGCTATGGCAAGGTGTGGAGCACTATGTGGGTGG 291
|
267 TCGCAACGGAGACCTGGAGGAGATCTTTTCACCCCTGGGGAATCAGTGATCCAGGTTTC 326
|
292 TCGCAACGGAGACCTGGAGGAGATCTTTTCACCCCTGGGGAATCAGTGATCCAGGTTTC 351
|
327 TGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGTATTTGTGACAGACAAAGGCCCGCTATCT 386
|
352 TGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGTATTTGTGACAGACAAAGGCCCGCTATCT 411
|
387 GTCTTTTGGGAAAGACAGTGGCAACAAGTTCAATGCCCTCCCTTGGACCCCAACACCGT 446
|
412 GTCTTTTGGGAAAGACAGTGGCAACAAGTTCAATGCCCTCCCTTGGACCCCAACACCGT 471
|
447 GTCTCCGCTTCATCAGTGGCGGCTGCTCTCTCATCGATGCCATTCGCTGCACTGGGA 506
|
472 GTCTCCGCTTCATCAGTGGCGGCTGCTCTCTCATCGATGCCATTCGCTGCACTGGGA 531
|
507 TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCTCCTCTC 547
|
532 TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCTCCTCTC 572
|
RESULT 8
AI991154/c 566 bp mRNA linear EST 27-OCT-1999
LOCUS wu38h12.x1 Soares_Dieckgraefe_colon_NHCD Homo sapiens cDNA clone
DEFINITION IMAGE:2522375 3' similar to TR:060844 060844 HOMOLOG OF RAT ZYMOGEN
GRANULE MEMBRANE PROTEIN. ; mRNA sequence.
ACCESSION AI991154
VERSION AI991154.1 GI:5838059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 566)
AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ccgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40up from Gibco
High quality sequence stop: 328.
FEATURES
source
Location/Qualifiers
1..566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2522375"
/clone_lib="Soares_Dieckgraefe_colon_NHCD"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site.1: Not 1; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGNAAGTGGGAGCGCGCTCTTTTTTTTTTTTTTTT 3'],
```

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library provided by Dr. Brian Dieckgraefe (Washington University, dieckelm.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 140 a 152 c 147 g 127 t

Query Match 49.7%; Score 497.8; DB 9; Length 566;
Best Local Similarity 96.5%; Pred. No. 2.6e-132;
Matches 520; Conservative 0; Mismatches 17; Indels 2; Gaps 1;

```

QY 73 GCCTCTGGCAATGCCATTCAGCCGAGGCTTCCTCTATAGTGGAGATGGAAGTGGT 132
Db 566 GCCTCTGGCAATGCCATTCAGCCGAGGCTTCCTCTATAGTGGAGATGGAAGTGGT 507
QY 133 GGTGGAAGCGGATTCCTCATCTCTGCAACCACTTGGAGCGGCCCATCACGCCCTCCGG 192
Db 506 GGTGGAAGCGGATTCCTCATCTCTGCAACCACTTGGAGCGGCCCATCACGCCCTCCGG 447
QY 193 GTCCAGTCAACACATACATCTAGTCTTTCAGGTGCGGTATGGCAAGTGTGGAGC 252
Db 446 GTCTGAGTCAACACATACATCTAGTCTTTCAGGTGCGGTATGGCAAGTGTGGAGC 387
QY 253 GACTATGTGGTGGTGGCAAGCGGAGCTTTCAGGTGCGGTATGGCAAGTGTGGAGC 312
Db 306 GACTATGTGGTGGTGGCAAGCGGAGCTTTCAGGTGCGGTATGGCAAGTGTGGAGC 327
QY 313 GTGATCCAGGTTTCTGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGG--TATTGTGACAG 370
Db 326 GTGATCCAGGTTTCTGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGGTTTGTGACCA 267
QY 371 ACAAGGGCCGTATCTGCTTTTGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGGTTTGTGACCA 267
Db 266 CAAGGGCCGTATCTGCTTTTGGGAAGTACAAAGTGGTACCTGGAAGAGCTGGGTTTGTGACCA 207
QY 431 TGCACCCCAACACCGTCTCGCTTCATCAGTGGCGGCTGCTCTCTCATCATGATGCCA 490
Db 206 TGCACCCCAACACCGTCTCGCTTCATCAGTGGCGGCTGCTCTCTCATCATGATGCCA 147
QY 491 TTGGCTGCTACTGGGATTTTACCCCACTAGCTGACAGATGCTGAGCCCTCTCTCTCT 550
Db 146 TTGGCTGCTACTGGGATTTTACCCCACTAGCTGACAGATGCTGAGCCCTCTCTCTCT 87
QY 551 GGCAGGGGCACTGTGATGAGGAGTAAGAACTCCCTTATCATCAACCCCAATCCAAATGG 609
Db 86 GGCAGGGGCACTGTGATGAGGAGTAAGAACTCCCTTATCATCAACCCCAATCCAAATGG 28

```

RESULT 9
B0028047/c
LOCUS 520 bp mRNA linear EST 27-MAR-2002
DEFINITION UI-H-CO0-arh-d-06-0-UI.sl NCI_CGAP_Sub9 Homo sapiens cDNA clone
IMAGE:3106570 3', mRNA sequence.
ACCESSION B0028047
VERSION B0028047.1 GI:19763326
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 520)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
UNPUBLISHED (1997)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgapbs-r@mail.nih.gov

FEATURES
source

Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers
1..520
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3106570"
/clone_lib="NCI_CGAP_Sub9"
/tissue_type="mixed"
/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p7T3-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
Cholonic mucosa with Crohns disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, Cervical
adenosquamous carcinoma, Ligament cells, prostate
carcinoma, Bladder carcinoma, Brain oligodendrocyte;
NCI_CGAP_Sub9 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into p7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are CGTC, AAGC, GGCC,
GGAAG, TAGC, TAAGC, ATGC, AGACA, ATCAC. For additional
information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_LIB=UI-H-CO0
TAG_TISSUE=colonic mucosa with Crohns disease
TAG_SEQ=CGTC"

BASE COUNT 119 a 134 c 136 g 131 t

ORIGIN

Query Match 46.9%; Score 469.4; DB 14; Length 520;
Best Local Similarity 99.8%; Pred. No. 3.8e-124;
Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 139 AAGCGATTCCTCATCTTGGCAACCACTTGGAGCGGCCCATCACCGCCCTCCGGTCCGA 198
Db 515 AGCGATTCCTCATCTTGGCAACCACTTGGAGCGGCCCATCACCGCCCTCCGGTCCGA 456
QY 199 GTCAACACATACATCATCTGAGTCTTTCAGTGGCTATGGCAAGTGTGGCGGACTAT 258
Db 455 GTCAACACATACATCATCTGAGTCTTTCAGTGGCTATGGCAAGTGTGGCGGACTAT 396
QY 259 GTGGTGGTCCGAACCGAGACCTGGAGGAGATCTTTTCACCCCTGGGGAATCAGTGATC 318
Db 395 GTGGTGGTCCGAACCGAGACCTGGAGGAGATCTTTTCACCCCTGGGGAATCAGTGATC 336
QY 319 CAGGTTCTGGGAAGTACAGTGGTACCTGAAAGCTGTTATTTGTGACAGACAAGGGC 378
Db 335 CAGGTTCTGGGAAGTACAGTGGTACCTGAAAGCTGTTATTTGTGACAGACAAGGGC 276
QY 379 CGCTATCTGCTTTTGGGAAAGACAGTGGCACAAGTTTCAATCCCTGCCCTTGGACCCC 438
Db 275 CGCTATCTGCTTTTGGGAAAGACAGTGGCACAAGTTTCAATCCCTGCCCTTGGACCCC 216
QY 439 AACACCGTGGTCCGCTTCATCAGTGGCGGCTGTGGTTCTCTCATCATGATGCCATG 498
Db 215 AACACCGTGGTCCGCTTCATCAGTGGCGGCTGTGGTTCTCTCATCATGATGCCATG 156
QY 499 CACTGGGATGTTTACCCCACTAGCTGACAGATGCTGAGCTCTCTCTCTCTCTCTCT 558
Db 155 CACTGGGATGTTTACCCCACTAGCTGACAGATGCTGAGCTCTCTCTCTCTCTCTCT 96

```

```

QY 559 CACTGTGATGAGGAGTAAAGAACTCCCTTATCACTAAACCCCATCAAAATGG 609
|||||
DB 95 CACTGTGATGAGGAGTAAAGAACTCCCTTATCACTAAACCCCATCAAAATGG 45

RESULT 10
BM987195/c
LOCUS
DEFINITION
  490 bp mRNA linear EST 25-MAR-2002
  UI-H-CO0-aql-a-10-0-UI-s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
  IMAGE: 3104154 3', mRNA sequence.
ACCESSION
  BM987195
VERSION
  BM987195.1 GI:19706584
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
  1 (bases 1 to 490)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. Jose Mercuende
  cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Cloning Distribution: Cloning Distribution Information can be found
  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source
  1..490
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE: 3104154"
  /clone_lib="NCI_CGAP_Sub9"
  /tissue_type="mixed"
  /dev_stage="mixed"
  /lab_host="DH10B (Life Technologies)"
  /note="Vector: p7T3-Pac (Pharmacia) with a modified
  polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
  Cholinic mucosa with Crohns disease, Cholinic mucosa with
  ulcerative colitis, Fetal thymus, Cervix, Cervical
  adenocarcinoma, Bladder carcinoma, Ligament cells, Prostate
  carcinoma, Bladder carcinoma, Brain oligodendrocyte;
  NCI_CGAP_Sub9 is a subtracted cDNA library constructed
  according to Bonaldo, Lennon and Soares, Genome Research,
  6:791-806, 1996. First strand cDNA synthesis was primed
  with an oligo-dT primer containing a Not I site. Double
  stranded cDNA was ligated to an EcoR I adaptor, digested
  with Not I, and cloned directionally into p7T3-Pac
  vector. The oligonucleotide used to prime the synthesis of
  first-strand cDNA contains a library tag sequence that is
  located between the Not I site and the (dT)18 tail. The
  sequence tags for this library are CGTC, AAGC, GGGCC,
  GGAAG, TAGC, TAAGC, ATGC, AGACA, ATCAC. For additional
  information, contact: Bento Soares, bento-soares@uiowa.edu
  TAG_LIB=UI-H-CO0
  TAG_TISSUE=colonic mucosa with Crohns disease
  TAG_SEQ=CGTC"
BASE COUNT 111 a 127 c 127 g 125 t
ORIGIN

Query Match 44.7%; Score 447; DB 14; Length 490;
Best Local Similarity 100.0%; Pred. No. le-117;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 CAGTTGAGCGCCCATCACCGCCCTCGGGTCCGAGTCACACATACATCATCGTAGGT 222
|||||
DB 490 CAGTTGAGCGCCCATCACCGCCCTCGGGTCCGAGTCACACATACATCATCGTAGGT 431

```


6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double-stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are CGTC, AACG, GGGCC, GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu

TAG_L1B=UI-H-COO
TAG_TISSUE=colonic mucosa with Crohns disease
TAG_SEQ=CGTC

BASE COUNT 110 a 125 c 126 g 124 t

Query Match 44.1%; Score 441; DB 14; Length 485;
Best Local Similarity 100.0%; Pred. No. 5.5e-116;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GACGGCCCATCACCGCCCTCCGGGTCCAGTCAACACATACATCGTAGTCTTCAG 228
|||||
Db 485 GACGGCCCATCACCGCCCTCCGGGTCCAGTCAACACATACATCGTAGTCTTCAG 426
|||||

QY 229 GTCCGCTATGGCAAGGTGGAGCGACTATGTGGTGGTCCGAACGGAGACCTGGAGGAG 288
|||||
Db 425 GTCCGCTATGGCAAGGTGGAGCGACTATGTGGTGGTCCGAACGGAGACCTGGAGGAG 366
|||||

QY 289 ATCTTTCTGCACCCCTGGGAATCAGTGATCCAGTTCTGGGAAGTACAAAGTGGTACCTG 348
|||||
Db 365 ATCTTTCTGCACCCCTGGGAATCAGTGATCCAGTTCTGGGAAGTACAAAGTGGTACCTG 306
|||||

QY 349 AAGAAGCTGGTATTTGTGACAGACAAGGCGCCTATCTGTCTTTTGGGAAGACAGTGGC 408
|||||
Db 305 AAGAAGCTGGTATTTGTGACAGACAAGGCGCCTATCTGTCTTTTGGGAAGACAGTGGC 246
|||||

QY 409 ACAAGTTTCAATGCGTCCCTTGGCACCAACACCGTGTCCCGTTCATCAGTGGCCGG 468
|||||
Db 245 ACAAGTTTCAATGCGTCCCTTGGCACCAACACCGTGTCCCGTTCATCAGTGGCCGG 186
|||||

QY 469 TCTGTTCTCTCATCGATGCCATTTGGCCTGCACTGGGATGTTTACCCCACTAGCTGCAGC 528
|||||
Db 185 TCTGTTCTCTCATCGATGCCATTTGGCCTGCACTGGGATGTTTACCCCACTAGCTGCAGC 126
|||||

QY 529 AGATCCTGAGCTCTCTCTTTGGCAGGGGCACTGTGATGAGGAGTAAGAATCCCTTTAT 588
|||||
Db 125 AGATCCTGAGCTCTCTCTTTGGCAGGGGCACTGTGATGAGGAGTAAGAATCCCTTTAT 66
|||||

QY 589 CACTAACCCCATCCAAATGG 609
|||||
Db 65 CACTAACCCCATCCAAATGG 45
|||||

RESULT 12
BM987303/c
LOCUS 492 bp mRNA linear EST 25-MAR-2002
DEFINITION UI-H-COO-aqp-e-10-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
IMAGE: 3104706 3', mRNA sequence.
ACCESSION BM987303
VERSION BM987303.1 GI:19706692
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 492)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
POLYA=Yes

FEATURES

Location/Qualifiers
1..492
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3104706"
/clone_lib="NCI_CGAP_Sub9"
/tissue_type="mixed"
/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; tissues: Colonic mucosa with Crohns disease, Colonic mucosa with ulcerative colitis, Fetal thymus, Cervix, Cervical adenosquamous carcinoma, Ligament cells, Prostate carcinoma, Bladder carcinoma, Brain oligodendroglia; NCI_CGAP_Sub9 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are CGTC, AACG, GGGCC, GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_L1B=UI-H-COO
TAG_TISSUE=colonic mucosa with Crohns disease
TAG_SEQ=CGTC

BASE COUNT 112 a 127 c 127 g 126 t

Query Match 43.5%; Score 435.4; DB 14; Length 492;
Best Local Similarity 99.8%; Pred. No. 2.3e-114;
Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 173 GCCCATCACCGCCCTCCGGTCCGAGTCAACACATACATCGTAGTCTTCAGGTGC 232
|||||
Db 481 GCCCATCACCGCCCTCCGGTCCGAGTCAACACATACATCGTAGTCTTCAGGTGC 422
|||||

QY 233 GCTATGGCAAGGTGGAGGCACTATGTGGTGGTCCGAACGGAGACCTGGAGGAGATCT 292
|||||
Db 421 GCTATGGCAAGGTGGAGGCACTATGTGGTGGTCCGAACGGAGACCTGGAGGAGATCT 362
|||||

QY 293 TTCTGCACCTGGGGAATCAGTATCCAGGTTCTTGGGAAGTACAACTGGTACCTGAAGA 352
|||||
Db 361 TTCTGCACCTGGGGAATCAGTATCCAGGTTCTTGGGAAGTACAACTGGTACCTGAAGA 302
|||||

QY 353 AGCTGGTATTTGTGACAGACAAGGCGCTATCTGCTTTTGGGAAGACAGTGGGCACAA 412
|||||
Db 301 AGCTGGTATTTGTGACAGACAAGGCGCTATCTGCTTTTGGGAAGACAGTGGGCACAA 242
|||||

QY 413 GTTTCATGCGCTCCCTTGCACCCCAACACCGTGTCTCCCGTTCATCAGTGGCGGCTCG 472
|||||
Db 241 GTTTCATGCGCTCCCTTGCACCCCAACACCGTGTCTCCCGTTCATCAGTGGCGGCTCG 182
|||||

QY 473 GTTTCATGCGATGCCATTTGGCTGAGTGTGTACCCCACTAGTGTGCGAGCAT 532
|||||
Db 181 GTTTCATGCGATGCCATTTGGCTGAGTGTGTACCCCACTAGTGTGCGAGCAT 122
|||||

QY 533 GCTGAGCCTCTCTCTTGGCAGGGCACTGTGATGAGGAGTAAGAATCCCTTATCACT 592
|||||
Db 121 GCTGAGCCTCTCTCTTGGCAGGGCACTGTGATGAGGAGTAAGAATCCCTTATCACT 62
|||||

QY 593 AACCCCATCCAAATGG 609
|||||

```

Db      61  AAAAAAAAAAAAAAAAAA
RESULT 13
AA828206/c
LOCUS   AA828206      497 bp      mRNA      linear      EST 20-FEB-1998
DEFINITION
of05d11.s1 NCI_CGAP_Col2 Homo sapiens cDNA clone IMAGE:1420245 3'
similar to TR:063680 063680 ZG-16P. ;, mRNA sequence.
ACCESSION
AA828206
VERSION  AA828206.1  GI:2900569
KEYWORDS EST.
SOURCE  human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 497)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@femail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 339.
FEATURES
Location/Qualifiers
source
1..497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1420245"
/clone_lib="NCI_CGAP_Col2"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Pooled colon tumors. 5' adaptor sequence: 5'
GAATTCGGCAGCAG 3' 3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 120 a 131 c 131 g 115 t
ORIGIN
Query Match 43.5%; Score 435.2; DB 9; Length 497;
Best Local Similarity 97.7%; Pred. No. 2.6e-114;
Matches 463; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
Qy 136 GGAAGCGATTCTCTATTCGTGGCAACAGTTGGACGGCCCATCACCCTCCGGGTC 195
Db |||||
497 GGAAGCGATTCTCTATTCGTGGCAACAGTT--GACGCCCATTCACCGGC--TCGGGTC 441
Qy 196 -CGAGTCAACACATACTACATCGTAGTCTTCAGGTGCGCTATGCAAGGTCTGGAGCGAC 255
Db |||||
440 CGAGTCAACACATACTACATCGTAGTCTTCAGGTGCGCTATGCAAGGTCTGGAGCGAC 381
Qy 256 TATGTGGTGGTCTGCAACGGAGACCTGGAGGAGATCTTCTGCAACCTGGGGGAATCAGTG 315
Db |||||
380 TATGTGGTGGTCTGCAACGGAGACCTGGAGGAGATCTTCTGCAACCTGGGGGAATCAGTG 321
Qy 316 ATCCAGGTCTTCTGGGAAGTCAAGTGTACTGAAGAAGCTGGTATTGTGACAGACAAG 375
Db |||||
320 ATCCAGGTCTTCTGGGAAGTCAAGTGTACTGAAGAAGCTGGTATTGTGACAGACAAG 261
Qy 376 GGCGGCTATCTGCTTTTGGGAAGACAGTGGCACAAGTTTCATGCGCTCCCTTCAC 435
Db |||||
260 GGCGGCTATCTGCTTTTGGGAAGACAGTGGCACAAGTTTCATGCGCTCCCTTCAC 201

```

```

Qy 436 CCCAACACCGCTCGCTTCATCAGTGGCGGCTGGTTCTCTCATCGATGCCATTGGC 495
Db |||||
200 CCCAACACCGCTCGCTTCATCAGTGGCGGCTGGTTCTCTCATCGATGCCATTGGC 141
Qy 496 CTCACATGGGATGTTTACCCCACTAGCTGAGCAGAGTGTGAGCGCTCTCTCTTGGCAG 555
Db |||||
140 CTCACATGGGATGTTTACCCCACTAGCTGAGCAGAGTGTGAGCGCTCTCTCTTGGCAG 81
Qy 556 GGCACATGTGATGAGGAGTAAGACACCTTATCACTAACCCCAATCAATGG 609
Db |||||
80 GGCACATGTGATGAGGAGTAAGAACTCCCTTATCACTAACCCCAATCAATGG 27

RESULT 14
LOCUS   AT1749030      455 bp      mRNA      linear      EST 22-JUN-1999
DEFINITION
ac38005.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:2374257 3' similar to TR:060844 060844 HOMOLOG OF RAT ZYMOGEN
GRANULE MEMBRANE PROTEIN. ;, mRNA sequence.
ACCESSION
AT1749030
VERSION  AT1749030.1  GI:5127229
KEYWORDS EST.
SOURCE  human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 455)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE   WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
source
1..455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2374257"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Organ: colon; Vector: pT7T3D-pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTAGCAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' AATTCACTAGTAAT 3' and 5' ATTACTAGT 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT7T3 vector. Library constructed by Bob
Barstead."
BASE COUNT 109 a 123 c 120 g 102 t 1 others
ORIGIN
Query Match 42.6%; Score 426.4; DB 9; Length 455;
Best Local Similarity 99.5%; Pred. No. 8.6e-112;
Matches 427; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 181 ACCGCCCTCCGGTCCGAGTCAACACATACATCTAGTCTTTCAGGTGCGGTATGGC 240
Db |||||
455 ACCGCCCTCCGGTCCGAGTCAACACATACATCTAGTCTTTCAGGTGCGGTATGGC 396

```

[illegible]

RESULT 15
AW001526/c
LOCUS
DEFINITION
AW001526 467 bp mRNA linear EST 09-MAR-2000
wu34a06.xl Soares_Dieckgraebe_colon_NHCD Homo sapiens cDNA clone
IMAGE:2521906 3' similar to TR:060844 O60844 HOMOLOG OF RAT ZYMOGEN
GRANULE MEMBRANE PROTEIN. ; mRNA sequence.

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 467)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

```
seq primer: -40UP from Gibco.
FEATURES
    source
        1..467
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:2521906"
                /clone_lib="Soares-Dieckgraefe_colon_NHCD"
                /tissue_type="colonic mucosa from 3 patients with Crohn's
```

BASE COUNT 112 a 125 c 125 g 105 t

Query Match	42.3%	Score 423.2	DB 10	Length 467
Best Local Similarity	99.1%	Pred. No. 7.3e-111		
Matches 436	Conservative 0	Mismatches 3	Indels 1	Gaps 1
171	CGGCCCCATACCGCCCTCCGGTCCGAGTCAACACATACATACATCGTAGGTCTTCAGGT	230		
172				
173				
174				
175				
176				
177				
178				
179				
180				
181				
182				
183				
184				
185				
186				
187				
188				
189				
190				
191	CGGTATGGCAAGGTGTGGACGCACTATGTGGGTGGTCGCAACGGAGACCTGGAGGAGAT	290		
192				
193				
194				
195				
196				
197				
198				
199				
200				
201	CGGTATGGCAAGGTGTGGACGCACTATGTGGGTGGTCGCAACGGAGACCTGGAGGAGAT	348		
202				
203				
204				
205				
206				
207				
208				
209				
210				
211				
212				
213				
214				
215				
216				
217				
218				
219				
220				
221				
222				
223				
224				
225				
226				
227				
228				
229				
230				
231				
232				
233				
234				
235				
236				
237				
238				
239				
240				
241				
242				
243				
244				
245				
246				
247				
248				
249				
250				
251				
252				
253				
254				
255				
256				
257				
258				
259				
260				
261				
262				
263				
264				
265				
266				
267				
268			</	

Search completed: July 16, 2003, 14:55:35
Job time : 1097.78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 10:35:27 ; Search time 39.0516 Seconds
(without alignments)
7860.974 Million cell updates/sec

Title: US-09-981-353-24
Perfect score: 1001
Sequence: 1 atgattacgaattcgagctc.....gtgagcaaaagccagcaaa 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2.6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545.8	54.5	560	1	US-08-469-667-10
2	545.8	54.5	560	4	US-09-224-110-10
3	545.8	54.5	560	5	PCT-US95-07289-10
4	352.2	35.2	3789	4	US-09-075-019-8
5	352	35.2	3983	4	US-09-481-049-1
6	351.8	35.1	4824	2	US-08-485-139-5
7	351.8	35.1	4824	3	US-08-750-357-5
8	351.8	35.1	8710	1	US-08-480-862B-3
9	351.8	35.1	8710	1	US-08-480-210-3
10	351.8	35.1	9019	1	US-08-480-882B-4
11	351.8	35.1	9019	1	US-08-480-210-4
12	350.8	35.0	738	1	US-08-726-462B-3
13	350.8	35.0	738	2	US-09-046-203-3
14	350.8	35.0	738	3	US-09-272-104-3
15	350.8	35.0	738	4	US-09-272-097-3
16	350.8	35.0	4283	1	US-08-343-401A-3
17	350.8	35.0	4283	1	US-08-445-265A-1
18	350.8	35.0	4283	3	US-08-990-442-1
19	350.8	35.0	10306	3	US-08-716-351A-4
20	350.8	35.0	10970	3	US-08-716-351A-5
21	350.6	35.0	5692	4	US-09-526-993-11
22	350.6	35.0	5737	4	US-09-526-993-9
23	350.6	35.0	6157	4	US-09-526-993-10
24	350.6	35.0	6202	4	US-09-526-993-8
25	350.2	35.0	716	4	US-08-998-416-55
26	350.2	35.0	752	4	US-08-976-259-108
27	350.2	35.0	2961	4	US-08-446-935-6

28	350.2	35.0	3699	3	US-08-646-538-6	Sequence 6, Appl
29	350.2	35.0	3699	4	US-09-503-222-6	Sequence 6, Appl
C 30	350.2	35.0	4145	3	US-08-651-472-62	Sequence 62, Appl
C 31	350.2	35.0	4145	4	US-08-358-928-62	Sequence 62, Appl
C 32	350.2	35.0	4277	3	US-08-651-472-63	Sequence 63, Appl
C 33	350.2	35.0	4277	4	US-08-358-928-63	Sequence 63, Appl
C 34	350.2	35.0	4435	2	US-08-792-824-1	Sequence 1, Appl
C 35	350.2	35.0	4525	2	US-08-613-861-2	Sequence 1, Appl
36	350.2	35.0	4539	1	US-08-119-512-1	Sequence 2, Appl
37	350.2	35.0	4539	1	US-08-488-015B-1	Sequence 1, Appl
38	350.2	35.0	4542	3	US-08-814-412-11	Sequence 11, Appl
C 39	350.2	35.0	4883	4	US-09-608-730B-21	Sequence 21, Appl
40	350.2	35.0	4965	3	US-08-675-566-22	Sequence 22, Appl
41	350.2	35.0	5033	1	US-08-038-768A-1	Sequence 1, Appl
42	350.2	35.0	5109	3	US-08-675-566-19	Sequence 19, Appl
43	350.2	35.0	5178	2	US-08-474-169-2	Sequence 2, Appl
44	350.2	35.0	5356	4	US-08-446-935-1	Sequence 1, Appl
45	350.2	35.0	5382	4	US-09-479-122-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-469-667-10
; Sequence 10, Application US/08469667
; Patent No. 5733748
; GENERAL INFORMATION:
; APPLICANT: YU, Guo-Liang
; TITLE OF INVENTION: Colon Specific Genes and Proteins
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: NJ
; COUNTRY: USA
; ZIP: 07068-1739
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,667
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..490
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 2..490
; US-08-469-667-10

Query Match 54.5%; Score 545.8; DB 1; Length 560;
Best Local Similarity 99.5%; Pred. No. 2.7e-158;

QY	661	TTGTTATTCGGCTCACAAATTCACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTG	720
Db	143	TTGTTATTCGGCTCACAAATTCACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTG	202
QY	721	GGGTGCCCTAATGAGTGAGCTAACTACACATTAATTGCGTTGCGCTACCTGCCCGCTTTCCA	780
Db	203	GGGTGCCCTAATGAGTGAGCTAACTACACATTAATTGCGTTGCGCTACCTGCCCGCTTTCCA	262
QY	781	GTCCGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGAGAGCGCG	840
Db	263	GTCCGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGAGAGCGCG	322
QY	841	TTTTCGGTATTGGCGCTCTCCGCTTCCTCGCTCACCTGACTCGCTCGCTCGCTTCG	900
Db	323	TTTTCGGTATTGGCGCTCTTCGGCTTCCTCGCTCACCTGACTCGCTCGCTCGCTTCG	382
QY	901	GCTCGGGGAGCGGTATCAGCTCACTCAAAGCGCGTAAATACGGTTATPCCACAGAATCAGG	960
Db	383	GCTCGGGGAGCGGTATCAGCTCACTCAAAGCGCGTAAATACGGTTATPCCACAGAATCAGG	442
QY	961	GGGATACCGAGGAAGAAGCATGTGAGCAAAAAGGCCAGCAA	1001
Db	443	GGATACCGAGGAAGAAGCATGTGAGCAAAAAGGCCAGCAA	483

RESULT 14
US-09-272-104--3
; Sequence 3, Application US/09272104
; Patent No. 6145434
; GENERAL INFORMATION:
; APPLICANT: Perkin-Elmer Corporation, Applied Biosystems
; APPLICANT: Division
; TITLE OF INVENTION: METHOD FOR DETECTING
; TITLE OF INVENTION: OLIGONUCLEOTIDES USING ENERGY TRANSFER DYES WITH LONG STROKE S
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; ADDRESSEE: & Rosati
; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Wordperfect for windows 6.0,
; SOFTWARE: ASCII (DOS) TEXT format
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,104
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/726,462
; FILING DATE: October 4, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/046,203
; FILING DATE: March 23, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: 16842-755
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 493-9300
; TELEFAX: (415) 493-6811

RESULT 15
US-09-272-097-3
; Sequence 3, Application US/09272097
; Patent No. 6335440
; GENERAL INFORMATION:
; APPLICANT: PE Corporation (NY)
; TITLE OF INVENTION: METHOD FOR DETECTING OLIGONUCLEOTIDES USING
; ENERGY TRANSFER DYES WITH LONG STROKE SHIFT
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; & Rosati
; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Word 97
; ASCII (DOS) TEXT format
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,097
; FILING DATE: 18-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 13:28:07 ; Search time 131.716 Seconds

(without alignments)
12133.023 Million cell updates/sec

Title: US-09-981-353-24

Perfect score: 1001

Sequence: 1 atgattacgaattcagctc.....gtgagcaaaaggccagcaaa 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1114047 seqs, 798260406 residues

Total number of hits satisfying chosen parameters: 2228094

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBOMB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	962	96.1	1001	9	US-09-981-353-24
2	583	58.2	737	9	US-10-235-994-17
3	582	58.1	736	9	US-10-158-646-5
C 4	352.8	35.2	7208	9	US-10-329-346-5
5	352.4	35.2	4069	9	US-10-251-385-287
C 6	352.4	35.2	4069	9	US-10-251-385-288
7	352	35.2	3983	10	US-09-758-987-1
C 8	351.8	35.1	8108	9	US-10-121-988-77
9	350.8	35.0	584	9	US-09-736-457-639
10	350.8	35.0	584	9	US-09-902-941-639
11	350.8	35.0	584	9	US-09-849-626-639
12	350.8	35.0	584	9	US-10-017-754-639
13	350.8	35.0	584	9	US-09-476-300-639
C 14	350.8	35.0	2571	9	US-10-137-036-93
15	350.8	35.0	3216	9	US-10-057-108-1
16	350.8	35.0	3357	9	US-10-161-403-96
17	350.8	35.0	3448	9	US-10-128-590-6
18	350.8	35.0	4118	10	US-09-068-821-17
19	350.8	35.0	4118	10	US-09-068-821-18

20	350.8	35.0	4162	10	US-09-861-881-1	Sequence 1, Appl1
21	350.8	35.0	4393	10	US-09-861-881-2	Sequence 2, Appl1
22	350.8	35.0	6688	9	US-09-813-453A-72	Sequence 72, Appl1
23	350.8	35.0	8320	9	US-09-813-453A-71	Sequence 71, Appl1
C 24	350.6	35.0	4689	9	US-10-057-108-5	Sequence 5, Appl1
C 25	350.6	35.0	6470	9	US-10-057-108-4	Sequence 4, Appl1
C 26	350.6	35.0	8330	9	US-10-057-108-3	Sequence 3, Appl1
C 27	350.2	35.0	752	10	US-09-956-004-108	Sequence 108, Appl1
C 28	350.2	35.0	1092	9	US-09-764-868-232	Sequence 232, Appl1
C 29	350.2	35.0	1092	9	US-09-764-868-232	Sequence 1853, Appl1
C 30	350.2	35.0	2074	9	US-10-106-698-350	Sequence 350, Appl1
C 31	350.2	35.0	2192	9	US-10-021-403A-10	Sequence 10, Appl1
C 32	350.2	35.0	3534	9	US-10-021-403A-9	Sequence 9, Appl1
C 33	350.2	35.0	3589	9	US-10-234-406-1	Sequence 1, Appl1
C 34	350.2	35.0	3609	9	US-10-234-406-3	Sequence 3, Appl1
C 35	350.2	35.0	3637	9	US-10-206-030-3	Sequence 3, Appl1
C 36	350.2	35.0	3637	9	US-10-206-030-4	Sequence 4, Appl1
C 37	350.2	35.0	3637	12	US-10-066-390-3	Sequence 3, Appl1
C 38	350.2	35.0	3637	12	US-10-066-390-4	Sequence 4, Appl1
C 39	350.2	35.0	3928	9	US-10-128-590-45	Sequence 45, Appl1
C 40	350.2	35.0	4119	10	US-09-993-170-1	Sequence 1, Appl1
C 41	350.2	35.0	4205	9	US-09-883-848A-35	Sequence 35, Appl1
C 42	350.2	35.0	4205	9	US-09-883-848A-48	Sequence 48, Appl1
C 43	350.2	35.0	4496	9	US-10-234-406-5	Sequence 5, Appl1
C 44	350.2	35.0	4704	9	US-09-932-328-4	Sequence 4, Appl1
C 45	350.2	35.0	4773	9	US-09-991-209-32	Sequence 32, Appl1

ALIGNMENTS

RESULT 1

US-09-981-353-24
; Sequence 24, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 3220207CB1
; NAME/KEY: unsure
; LOCATION: 610-648
; OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-24

Qy	1	ATGATTACGAATTCGAGCTCGTACCGCCCGAGGTCCTTCCTCTCTAGTGGAGAG	120
Db	1	ATGATTACGAATTCGAGCTCGTACCGCCCGAGGTCCTTCCTCTCTAGTGGAGAG	60
Qy	61	CTCTGTGCTCAGCCTCTGCAATGCCATTCAGCCAGGTCCTTCCTCTCTAGTGGAGAG	120
Db	61	CTCTGTGCTCAGCCTCTGCAATGCCATTCAGCCAGGTCCTTCCTCTCTAGTGGAGAG	120
Qy	121	TATGGAAGTGGTGGAAAGCGATTCTCTATTCTGGCAACAGTTGGAGCGGCCCATC	180
Db	121	TATGGAAGTGGTGGAAAGCGATTCTCTATTCTGGCAACAGTTGGAGCGGCCCATC	180
Qy	181	ACGCGCCCTCGGGTCCGAGTCAACATACATCTAGTGTCTTTCAGTSCGCTATGGC	240

```

181 ACCGCGCTCCGGTCCGAGTCAACACATACATCGTAGGCTTCAGGTGGCTATGCG 240
241 AAGGTGTGAGGAGTATGTGGTGGTCCGACGAGACCTGGAGGAGATCTTTCTGCAC 300
241 AAGGTGTGAGGAGTATGTGGTGGTCCGACGAGACCTGGAGGAGATCTTTCTGCAC 300
301 CTTGGGGAATCAGTATCAGTCCAGTTTCTGGGAAGTACAAGTGTACCTGAAGAAGCTGGTA 360
301 CTTGGGGAATCAGTATCAGTCCAGTTTCTGGGAAGTACAAGTGTACCTGAAGAAGCTGGTA 360
361 TTGTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAAGTTTCAAT 420
361 TTGTGACAGACAAGGGCGCTATCTGTCTTTTGGGAAGACAGTGGCACAAAGTTTCAAT 420
421 GCGTCCCGTTCGACCCCAACACGCTGCTCCGCTTCATCAGTGGCGGTCTGGTCTCTC 480
421 GCGTCCCGTTCGACCCCAACACGCTGCTCCGCTTCATCAGTGGCGGTCTGGTCTCTC 480
481 ATCGATGCGATGGCTGCTGCTGGGATGTTTACCCACTAGCTGCAGCAGATGCTGAGCC 540
481 ATCGATGCGATGGCTGCTGCTGGGATGTTTACCCACTAGCTGCAGCAGATGCTGAGCC 540
541 TCCTCTCCTTGGCAGGGGCACTGTGATGAGGAGTAAAGTCCCTTATCACTAACCCCA 600
541 TCCTCTCCTTGGCAGGGGCACTGTGATGAGGAGTAAAGTCCCTTATCACTAACCCCA 600
601 TCCAAATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
601 TCCAAATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
661 TTGTTATCGCTCACAATTCACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTG 720
661 TTGTTATCGCTCACAATTCACACAACATACGAGCGGGAAGCATAAAGTGTAAAGCCTG 720
721 GGGTGCCTAATGAGTGAGTAACTCACTCACTAATTAATGCGTTCGCTCACTGCCGCTTCCA 780
721 GGGTGCCTAATGAGTGAGTAACTCACTCACTAATTAATGCGTTCGCTCACTGCCGCTTCCA 780
781 GTCGGGAACCTGCTGCGACGCTCATTAATGAATCGGCAACGCGGGGAGAGCGG 840
781 GTCGGGAACCTGCTGCGACGCTCATTAATGAATCGGCAACGCGGGGAGAGCGG 840
841 TTGCGTATTTGGCGCTCTTCGCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTG 900
841 TTGCGTATTTGGCGCTCTTCGCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTG 900
901 GCTGCGGAGCGGTATCAGCTCACTCAAGCGGTATACGGTTATCCACAGAATCAGG 960
901 GCTGCGGAGCGGTATCAGCTCACTCAAGCGGTATACGGTTATCCACAGAATCAGG 960
961 GGGATACGAGGAAAGACATGTGAGCAAAAGGCCAGCAAA 1001
961 GGGATACGAGGAAAGACATGTGAGCAAAAGGCCAGCAAA 1001
```

```

RESULT 2
US-10-235-994-17
; Sequence 17, Application US/10235994
; Publication No. US2003010102A1
; GENERAL INFORMATION:
; APPLICANT: Bartha, Gabor
; APPLICANT: Walker, Michael
; TITLE OF INVENTION: METHODS FOR ANALYZING GENE EXPRESSION PATTERNS
; FILE REFERENCE: ICYTP012
; CURRENT APPLICATION NUMBER: US/10/235,994
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US/10/003,608
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/245,081
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 17
; LENGTH: 737
; TYPE: DNA
; ORGANISM: Human
US-10-235-994-17

Query Match
Best Local Similarity 58.2%; Score 583; DB 9; Length 737;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CCCAAGAATGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCTCTGCGCAATGC 86
DB 129 CCCAAGAATGTGACAGTCGCTCTCTAGCCCTTCTGTGCTCAGCTCTGCGCAATGC 188
QY 87 CAATCAGCCAGGCTTCTCTCTATAGTGGAGATGGAAGTGTGGTGGAAAGCGATT 146
DB 189 CAATCAGCCAGGCTTCTCTCTATAGTGGAGATGGAAGTGTGGTGGAAAGCGATT 248
QY 147 CTCTCATTTCTGCAACACAGTTGGAGCGGCCCATCATACCCGCTTCCGGGTCGAGTCAACAC 206
DB 249 CTCTCATTTCTGCAACACAGTTGGAGCGGCCCATCATACCCGCTTCCGGGTCGAGTCAACAC 308
QY 207 ATACTACATCTGATGCTTTCAGGTGCTGATGGCAAGGTGGAGCGACTATGTGGGTGG 266
DB 309 ATACTACATCTGATGCTTTCAGGTGCTGATGGCAAGGTGGAGCGACTATGTGGGTGG 368
QY 267 TCGCAACGAGACCTGGAGGAGATCTTTCTGCAACCTCGGGAATCAGTATCCAGGTTTC 326
DB 369 TCGCAACGAGACCTGGAGGAGATCTTTCTGCAACCTCGGGAATCAGTATCCAGGTTTC 428
QY 327 TGGGAAGTACAAGTGGTACCTGGAAGTGGTATTTGTGACAGACAAGGGCCGCTATCT 386
DB 429 TGGGAAGTACAAGTGGTACCTGGAAGTGGTATTTGTGACAGACAAGGGCCGCTATCT 488
QY 387 GTCCTTTGGGAAGACAGTGGCACAAGTTTCAATGCGCTCCCTTGCACCCCAACACCGT 446
DB 489 GTCCTTTGGGAAGACAGTGGCACAAGTTTCAATGCGCTCCCTTGCACCCCAACACCGT 548
QY 447 GTCGCGCTTCATCAGTGGCGGCTGCTGCTCTCATGATGCAATGGCCTGCACCTGGGA 506
DB 549 GTCGCGCTTCATCAGTGGCGGCTGCTGCTCTCATGATGCAATGGCCTGCACCTGGGA 608
QY 507 TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGCACTGTGA 566
DB 609 TGTTTACCCCACTAGCTGCAGCAGATGCTGAGCCCTCTCTCTTGGCAGGGGCACTGTGA 668
QY 567 TGAGGAGTAAGAATCCCTTATCACTAACCCCAATCCAAATGG 609
DB 669 TGAGGAGTAAGAATCCCTTATCACTAACCCCAATCCAAATGG 711
```

```

RESULT 3
US-10-158-646-5
; Sequence 5, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Sornasse, Thierry
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030073105A1 239588.4
```


US-10-158-646-5

Query Match 58.18; Score 582; DB 9; Length 736;
Best Local Similarity 100.0%; Pred. No. 5.8e-183;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CCCAGAAATGTTGACAGTCTCTCTAGCCCTTCTGTGCTCAGCCCTCTGCGCAATGC 86
DB 129 CCCAGAAATGTTGACAGTCTCTCTAGCCCTTCTGTGCTCAGCCCTCTGCGCAATGC 188
QY 87 CATTAGCCAGCTCTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGAAAGCGATT 146
DB 189 CATTAGCCAGCTCTCTCTCTATAGTGGAGAGTATGGAAGTGGTGGAAAGCGATT 248
QY 147 CTCTCATTTCTGGCAACAGTTGACGCGCCCATCAGCCCTCCGGTCCGGTCCAGTCAACAC 206
DB 249 CTCTCATTTCTGGCAACAGTTGACGCGCCCATCAGCCCTCCGGTCCGGTCCAGTCAACAC 308
QY 207 ATACTACATCGTAGGTCTTACAGTGGCGCTATGGCAAGTGTGGAGGACTATGTGGGTGG 266
DB 309 ATACTACATCGTAGGTCTTACAGTGGCGCTATGGCAAGTGTGGAGGACTATGTGGGTGG 368
QY 267 TCGCAACGAGGACCTCGAGGAGATCTTCTGCACCCCTGGGGAATCAGTATCCAGTTTC 326
DB 369 TCGCAACGAGGACCTCGAGGAGATCTTCTGCACCCCTGGGGAATCAGTATCCAGTTTC 428
QY 327 TGGGAAGTACAAGTGTGACTGAAGAGCTGGTATTTGTGACAGCAAGGGCCGCTATCT 386
DB 429 TGGGAAGTACAAGTGTGACTGAAGAGCTGGTATTTGTGACAGCAAGGGCCGCTATCT 488
QY 387 GTCTTTTGGGAAGACAGTGGCAAGTTTCAATGCCCTCCCTTGCACCCCAACACCGT 446
DB 489 GTCTTTTGGGAAGACAGTGGCAAGTTTCAATGCCCTCCCTTGCACCCCAACACCGT 548
QY 447 GCTCCGCTTCATCAGTGGCGGCTGCTGCTCTCTCATGATGCCATTGGCTGCATGGGA 506
DB 549 GCTCCGCTTCATCAGTGGCGGCTGCTGCTCTCTCATGATGCCATTGGCTGCATGGGA 608
QY 507 TGTTTACCCCACTAGCTGCAGAGATGCTGAGCCCTCTCTCTGCGAGGGCACTGTGA 566
DB 609 TGTTTACCCCACTAGCTGCAGAGATGCTGAGCCCTCTCTCTGCGAGGGCACTGTGA 668
QY 567 TGAGGAGTAAGAACTCCCTTATCATAACCCCACTCAAAATG 608
DB 669 TGAGGAGTAAGAACTCCCTTATCATAACCCCACTCAAAATG 710

RESULT 4

US-10-229-346-5/c
; Sequence 5, Application US/10229346
; Publication No. US20030120054A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Eric
; APPLICANT: Stacy, Cheryl
; TITLE OF INVENTION: Modified Cry3A Toxins
; FILE REFERENCE: 60065A
; CURRENT APPLICATION NUMBER: US/10/229,346
; CURRENT FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/316,421
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 7208
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: pCIB6850
US-10-229-346-5

Query Match 35.2%; Score 352.8; DB 9; Length 7208;
Best Local Similarity 82.4%; Pred. No. 2.4e-106;

Matches 378; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 543 CTCCTCTGGCAGGGCAGCTGTGATGAGAGTAAGAATCCCTTATCATCAACCCCAATC 602
DB 4727 CTATACCTTACAGCCCGTCGCAAGATATTGTTGGATCGTTGCTGCTCCATCCCTTAC 4668
QY 603 CAAATGNN 662
DB 4667 AATTGTTGTCATGTGCAAGCTTGGCGTATCATGTGCTAGCTGTTTCTGTGGAATTT 4608
QY 663 GTTATCCGCTCACAATTCACACAATACGAGCGGAGCAATAAAGTGTAAAGCCTGG 722
DB 4607 GTTATCCGCTCACAATTCACACAATACGAGCGGAGCAATAAAGTGTAAAGCCTGG 4548
QY 723 GTGCTTAATGAGTGAAGTAACTACATTAATTTGCGTTCGCTCACTGCCCTTCCAGT 782
DB 4547 GTGCTTAATGAGTGAAGTAACTACATTAATTTGCGTTCGCTCACTGCCCTTCCAGT 4488
QY 783 CGGGAACCTGTCGTGCGAGTGCATTAATGAATGGCCAAACGCGCGGAGAGCGGTT 842
DB 4487 CGGGAACCTGTCGTGCGAGTGCATTAATGAATGGCCAAACGCGCGGAGAGCGGTT 4428
QY 843 TCGCTATTGGGCGCTCTTCCGCTTCCCTCACTCAAAAGCGGTAATACGCTTATCCACAGAATCAGGG 902
DB 4427 TCGCTATTGGGCGCTCTTCCGCTTCCCTCACTCAAAAGCGGTAATACGCTTATCCACAGAATCAGGG 4368
QY 903 TCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGCTTATCCACAGAATCAGGG 962
DB 4367 TCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGCTTATCCACAGAATCAGGG 4308
QY 963 GATAGCGAGGAAGAACATGTGAGCAAAAGCCAGCAAA 1001
DB 4307 ATAAGCGAGGAAGAACATGTGAGCAAAAGCCAGCAAA 4269

RESULT 5

US-10-251-385-287
; Sequence 287, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 287
; LENGTH: 4069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-287

Query Match 35.2%; Score 352.4; DB 9; Length 4069;
Best Local Similarity 88.8%; Pred. No. 2.6e-106;
Matches 356; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 601 TCCAAATGNN 660
DB 1299 TCCAAAGCTCCCTCGAGAGCTTGGCGTAATCATGCTATAGCTGTTCTCTGTGAAA 1358
QY 661 TTGTTATCCGCTCACAATTCACACAATACGAGCGGAGCAATAAAGTGTAAAGCCTG 720
DB 1359 TTGTTATCCGCTCACAATTCACACAATACGAGCGGAGCAATAAAGTGTAAAGCCTG 1418
QY 721 GGTGCTTAATGAGTGAAGTAACTACATTAATTTGCGTTGCGCTCACTGCCCTTCCA 780

RESULT.T 1A

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 03:49:46 ; Search time 2395.68 Seconds
(without alignments)
15998.962 Million cell updates/sec

Title: US-09-981-353-81
Perfect score: 1317
Sequence: 1 aaggaacaaagtaagtccat.....ttttccttaataaaatgctc 1317

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_ph.*
- 7: gb_pl.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vl.*
- 14: em_ba.*
- 15: em_fun.*
- 16: em_hum.*
- 17: em_in.*
- 18: em_mu.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pi.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_un.*
- 28: em_vl.*
- 29: em_htg_hum.*
- 30: em_htg_inv.*
- 31: em_htg_other.*
- 32: em_htg_mus.*
- 33: em_htg_pln.*
- 34: em_htg_rtd.*
- 35: em_htg_mam.*
- 36: em_htg_vrt.*
- 37: em_sy.*
- 38: em_hgo_hum.*
- 39: em_hgo_mus.*
- 40: em_hgo_other.*
- 41: em_hgo_hum.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1109.2	84.2	1182	9	AK000224	BC000224 Homo sapi
2	1108	84.1	1176	9	BC029793	BC029793 Homo sapi
C 3	551	41.8	1222	4	SSR236932	AJ236932 Sus scrofa
C 4	399	30.3	166793	2	AC090401	AC090401 Homo sapi
5	399	30.3	167934	9	AP001034	AP001034 Homo sapi
6	303.6	23.1	351	6	AX245564	AX245564 Sequence
7	292.6	22.2	342	6	AX245708	AX245708 Sequence
8	291.2	22.1	373	6	AX071292	AX071292 Sequence
9	288.6	21.9	383	6	AX245551	AX245551 Sequence
10	276	21.0	370	6	AX071281	AX071281 Sequence
11	81.2	6.2	852	9	AF350504	AF350504 Homo sapi
12	81.2	6.2	1305	9	BC022895	BC022895 Homo sapi
13	81.2	6.2	1336	6	AX368966	AX368966 Sequence
14	81.2	6.2	1369	9	AF237905	AF237905 Homo sapi
15	79.6	6.0	861	6	AX369167	AX369167 Sequence
16	79.6	6.0	1155	6	AX369163	AX369163 Sequence
17	79.6	6.0	1353	6	AX369163	AX369163 Sequence
18	75.6	5.7	950	10	AF321128	AF321128 Mus muscu
19	75.4	5.7	518	10	AF280401	AF280401 Mus muscu
20	74.8	5.7	862	9	AB013104	AB013104 Homo sapi
21	74.8	5.7	930	9	AF286866	AF286866 Homo sapi
22	74.8	5.7	932	6	AX357469	AX357469 Sequence
23	74.8	5.7	981	9	AF354930	AF354930 Homo sapi
24	74.8	5.7	985	9	AF350502	AF350502 Homo sapi
25	74.8	5.7	1036	6	AX357467	AX357467 Sequence
26	74.8	5.7	1131	9	AF212240	AF212240 Homo sapi
27	74.8	5.7	1152	9	AF237908	AF237908 Homo sapi
28	74.8	5.7	1258	9	AF350503	AF350503 Homo sapi
29	74.8	5.7	1289	9	AF142409	AF142409 Homo sapi
30	74.8	5.7	1433	9	BC022854	BC022854 Homo sapi
31	74.8	5.7	1488	6	AX017505	AX017505 Sequence
32	74.8	5.7	1679	6	AX017610	AX017610 Sequence
33	71.2	5.4	2542	10	RATFCIGA	M22923 Rat Ige Pc
34	71.2	5.4	2545	6	ARI23798	ARI23798 Sequence
35	71.2	5.4	2545	6	I09687	I09687 Sequence 5
36	70.6	5.4	1025	10	AF237906	AF237906 Mus muscu
37	70.6	5.4	1557	9	BC031610	BC031610 Homo sapi
38	69.2	5.3	708	6	AX101306	AX101306 Sequence
39	69.2	5.3	1010	10	MUSFCERB	J05019 Mouse mast
C 40	68.6	5.2	107211	2	AC111330	AC111330 Rattus no
41	67.4	5.1	1066	10	AB026044	AB026044 Mus muscu
42	61.6	4.7	2292	9	AB071108	AB071108 Macaca fa
43	60.4	4.6	1018	4	OAR318333	AJ318333 Ovis arie
C 44	60	4.6	489	6	AX341797	AX341797 Sequence
45	60	4.6	894	9	HUMIGERB	D10583 Homo sapien

ALIGNMENTS

RESULT 1
AK000224
LOCUS Homo sapiens cDNA FLJ20217 fis, clone COLF3334.
DEFINITION AK000224
ACCESSION AK000224
VERSION AK000224.1 GI:7020167
KEYWORDS Oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone COLF3334.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Watanabe, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Pred. No. is the number of results predicted by chance to have a

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
2 (bases 1 to 1182)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp,
tel: 81-3-5449-5286, fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction; 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source

Location/Qualifiers
1..1182
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COLF3334"
/tissue_type="colon mucosa"
/clone_lib="Colf"
/note="cloning vector pME18SFL3"
1..1182
/gene="MS4A12"
58..861
/gene="MS4A12"
/codon_start=1
/product="MS4A12 protein"
/protein_id="BAA91020.1"
/db_xref="GI:7020168"
/translation="MMSKPTSHAEVNETIPNPPGSPFMAPGQFQPLGSLNLENQAO
GAQRAPIGTSFPGIFASVPGQGNQINPVSQAVNKEAKALGVQIHWGLMH
IGFIVICLISFSPGIVFASVIGVYFPGWGLSFIISGSLSVSKSLRCLVKG
SLGNIVSSILAFVIGVILVDMCINGVAGDYGWVLSGKISATLIMFSLLEFFVAC
ATAHFANQANTTNNMVLVPIPNMYESPVPASSAPPRCNNYSANAPK"
BASE COUNT 344 a 252 c 245 g 341 t
ORIGIN

Query Match 84.2%; Score 1109.2; DB 9; Length 1182;
Best Local Similarity 98.0%; Pred. No. 9.1e-270;
Matches 1123; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 172 TGAACGAACAAAGAAACAAACAGAGAGGAATGAAACACACATAATGATGTCATC 231
DB 9 TGGAGCAGAGAAAGAGAAACATAGAGTGCCAAAGGAACAAAGACATAATGATGTCATC 68
QY 232 CAAGCCCAACAGCCATGCTGAAGTAAATGAAACCATACCCAAACCCCTTACCCCAACAGCAG 291
DB 69 CAAGCCCAACAGCCATGCTGAAGTAAATGAAACCATACCCAAACCCCTTACCCCAACAGCAG 128
QY 292 CTTTATGGCTCCTGGATTCAACAGCCTCTGGGTTCAATCAACTTAGAAACCAAGCTCA 351
DB 129 CTTTATGGCTCCTGGATTCAACAGCCTCTGGGTTCAATCAACTTAGAAACCAAGCTCA 188
QY 352 GGGTGTCTCAGGCTCTCAGCCCTACGGCATCACATCTCCGGGAATCTTGTCTACAGTCA 411
DB 189 GGGTGTCTCAGGCTCTCAGCCCTACGGCATCACATCTCCGGGAATCTTGTCTACAGTCA 248
QY 412 ACCGGGTCAAGGAATATACAAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTT 471
DB 249 ACCGGGTCAAGGAATATACAAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTT 308
QY 472 TAAAGAGAACCAAGGACTAGGGGTGATCCAGATCATGTTGCGATTGATGCACATTGG 531
DB 309 TAAAGAGAACCAAGGACTAGGGGTGATCCAGATCATGTTGCGATTGATGCACATTGG 368
QY 532 TTTTGGGAATCTTTTGTGTTTAAATATCTCTCTTTTAGAGAACTATTAGTTTGGCTC 591
DB 369 TTTTGGGAATCTTTTGTGTTTAAATATCTCTCTTTTAGAGAACTATTAGTTTGGCTC 428

QY 592 TACTGCTGTTATTGTTGGATACCACTTCTGGGGTGGCCTTTCTTTTATATCTCTGGCTC 651
DB 429 TACTGCTGTTATTGTTGGATACCACTTCTGGGGTGGCCTTTCTTTTATATCTCTGGCTC 488
QY 652 TCTCTCTGTGTGTCAGCATCCAAAGAGCTTTCCGTTGTCTGTGTGTAAGGAGCCTTGGGAAT 711
DB 489 TCTCTCTGTGTGTCAGCATCCAAAGAGCTTTCCGTTGTCTGTGTGTAAGGAGCCTTGGGAAT 548
QY 712 GAACATGTTAGTTTCTATCTTTGGCCTTTCATTGGAGTGATTTCTGCTGCTGGTGGATATGTG 771
DB 549 GAACATGTTAGTTTCTATCTTTGGCCTTTCATTGGAGTGATTTCTGCTGCTGGTGGATATGTG 608
QY 772 CATCAATGGGGTAGCTGGCCAAAGACTACTGGGGCGTCTTTCTTGAAAGAGCATTTCAGC 831
DB 609 CATCAATGGGGTAGCTGGCCAAAGACTACTGGGGCGTCTTTCTTGAAAGAGCATTTCAGC 668
QY 832 CACGCTGATGATCTTCTCCCTCTTTGGAGTCTTCTGAGTGTGTCACAGCCCATTTTGC 891
DB 669 CACGCTGATGATCTTCTCCCTCTTTGGAGTCTTCTGAGTGTGTCACAGCCCATTTTGC 728
QY 892 CAACCAAGCAACACCAACCAATATGCTGCTCTGTTATTCCTGTTATTCCTGTTATGTTGAAAG 951
DB 729 CAACCAAGCAACACCAACCAATATGCTGCTCTGTTATTCCTGTTATTCCTGTTATGTTGAAAG 788
QY 952 CAACCCCTGTGACACAGCGTCTTCTTTCAGCTCTCCAGATGTCACCAACTACTCAGCTAA 1011
DB 789 CAACCCCTGTGACACAGCGTCTTCTTTCAGCTCTCCAGATGTCACCAACTACTCAGCTAA 848
QY 1012 TGCCCCCTAAATAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGAGAAAACACTACTT 1071
DB 849 TGCCCCCTAAATAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGAGAAAACACTACTT 908
QY 1072 GCAAAACTCTCTTAAAGAGATGCTCTTTATTTGCTACAATGATTTCTAGTCTTTTAAAC 1131
DB 909 GCAAAACTCTCTTAAAGAGATGCTCTTTATTTGCTACAATGATTTCTAGTCTTTTAAAC 968
QY 1132 TGTGTTTGTGAGATTTGTTTATAGTTGGTGGCTAATGATGGCTGTATCTCCCTCAGTCTC 1191
DB 969 TGTGTTTGTGAGATTTGTTTATAGTTGGTGGCTAATGATGGCTGTATCTCCCTCAGTCTC 1028
QY 1192 TCTTCTCATATTACCACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTAAAAAT 1251
DB 1029 TCTTCTCATATTACCACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTAAAAAT 1088
QY 1252 GATTCTGCAACTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTACTTTTCTCTTAATAA 1311
DB 1089 GATTCTGCAACTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTACTTTTCTCTTAATAA 1148
QY 1312 AATGTC 1317
DB 1149 AATGTC 1154

RESULT 2

BC029793
LOCUS
DEFINITION
BC029793 Homo sapiens, membrane-spanning 4-domains, subfamily A, member 12
4-domains, subfamily A, member 7, clone MGC:34288 IMAGE:5185424,
mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BC029793
BC029793.1 GI:20988642
MGC.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1176)
Strausberg, R.
Direct Submission
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

Db	405	AGTCCCTTGACGACAGATTGTCCTCAGTTCTCCCCCACTTCCGGCCACTTACGTCACA	346
Qy	1004	TCAGCTAAATGCCCTTAATAGTAAAGAGAAAGGGGTATCAGTCTTAATCTCATGTGAGAAA	1063
Db	345	CCACCTATACCACTAGGTATTAAAGAGAAAGCAATCCTGCATTTTCATGGGGTAG	286
Qy	1064	AACACT-TGCAGAAACTTCTTAAAGAGATGTCTTTTATTGTCTACAATGATTTCTAGTC	1122
Db	285	AACCGCTCTGCAGAAACTTCTTAAAGAGTTGGTCTTCGCGCTTCACGATGATTGCTAATC	226
Qy	1123	TTTAAAGACTGCTTTGAGATTTGTTTAAAGTTAGTTGGTCGTATGATGGCTGTATCTCCC	1182
Db	225	TTTAAAGACTGCTTTGAGATTTGTTTAAAGTTAGTTGGTCGTATGATGGCTGTATCTCCC	166
Qy	1183	T-----TCACTGCTCTTCTTACATTACCACTACTACATGCTGCGAAAGTGAAGGATCA	1237
Db	165	TCTCTACCACTGTTTCTTCTCTACA-TGCCACTACCACTGCTGCGAAAGTGAAGGCTG	107
Qy	1238	GAGGACTGAAAAATGATTTCTGCAACTCTCTTAAAGTTAGAAATGTTTCTGTCATATTAC	1297
Db	106	GAGGACTG-CARATGACTCTGCAACTCTCTTAAAGTTAGAAAT-TTCCCTGTTAAAGTTAC	49
Qy	1298	TTTTTCTCTTAATAAATGT	1316
Db	48	CCTTCCCTCAATAAATGT	30
RESULT 4			
AC090401/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

QY 1159 TCGCTAATGATGGGTGTTATCTCCCTTCACTGCTCTTCTTCTACATTACCACTACTACATGTC 1218
Db 100300 TCGCTAATGATGGGTGTTATCTCCCTTCACTGCTCTTCTTCTACATTACCACTACTACATGTC 100359
QY 1219 TGGCAAGGTGAGGATGAGGAGTGAAGAAATGTTCTGCAACTCTCTTAAAGTTAGAA 1278
Db 100360 TGGCAAGGTGAGGATGAGGAGTGAAGAAATGTTCTGCAACTCTCTTAAAGTTAGAA 100419
QY 1279 ATGTTCTGTCATATTTCTTCTTCTTAAATAAATGTC 1317
Db 100420 ATGTTCTGTCATATTTCTTCTTCTTAAATAAATGTC 100458
RESULT 6
LOCUS AX245564 351 bp DNA linear PAT 28-SEP-2001
DEFINITION Sequence 494 from Patent WO0166753.
ACCESSION AX245564
VERSION AX245564.1 GI:15860238
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D.,
Sudduth-Klinger, J., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Kassam, A., Lamson, G., Drmanac, R., Crkvenjakov, R., Dickson, M.,
Drmanac, S., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and
Stache-Crain, B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0166753-A 494 13-SEP-2001;
Chiron Corporation (US); Hyseq Inc. (US)
FEATURES
source
1. 351
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 123 a 81 g 68 t
ORIGIN
Query Match 23.1%; Score 303.6; DB 6; Length 351;
Best Local Similarity 93.0%; Pred. No. 5.7e-66;
Matches 318; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 172 TGAACGAACAAAGAAAGAAACAGAGAGGAATGAAAGACATATGATGTCATC 231
Db 10 TGGAGCAGAGAAAGAGGAAACATAGAGGTGCCAAGGAACAAAGACATATGATGTCATC 69
QY 232 CAAGCCACAGCCATGCTGAAGTAATGAACCATACCCAAACCCCTTACCACCAAGCAG 291
Db 70 CAAGCCACAGCCATGCTGAAGTAATGAACCATACCCAAACCCCTTACCACCAAGCAG 129
QY 292 CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCATCACTAGAAAACCAAGCTCA 351
Db 130 CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCATCACTAGAAAACCAAGCTCA 189
QY 352 GGGTGCTCAGGCTGCTCAGCCCTAGCGCATCACATCTCCGGGAATCTTTGGTAGCAGTCA 411
Db 190 GGGTGCTCAGGCTGCTCAGCCCTAGCGCATCACATCTCCGGGAATCTTTGGTAGCAGTCA 249
QY 412 ACCGGGTCAAGGAATATACAAATGATAAATCAAGTGTGGGAACAGCAGTAATGAAGTT 471
Db 250 ACCGGGTCAAGGAATATATATAATGATAAATCAAGTGTGGGAACAGCAGTAATGAAGTT 309
QY 472 TAAAGAACAAAGGCACTAGGGGTGATCC 503
Db 310 TAAAGAACAAAGGCACTAGGGGTGATCC 341
RESULT 8
LOCUS AX071292 373 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1764 from Patent WO0102568.
ACCESSION AX071292
VERSION AX071292.1 GI:12581643
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lamson, G., Drmanac, R., Crkvenjakov, R., Drmanac, S., Dickson, M.,
Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1764 11-JAN-2001;
CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
source
1. 373
/organism="Homo sapiens"
/db_xref="taxon:9606"

AF350504
LOCUS Homo sapiens four-span transmembrane protein 4 (4SPAN4) mRNA, complete cds.
ACCESSION AF350504
VERSION AF350504.1 GI:18028935
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS Hulett,M.D., Pagler,E., Hogarth,M., Eyre,H., Baker,E., Crawford,J., Sutherland,G.R. and Parish,C.R.
TITLE Isolation of a family of hematopoietic-expressed four-transmembrane genes related to CD20 and Fc epsilonRI-beta
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 852)
AUTHORS Hulett,M.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2001) Immunology and Cell Biology, John Curtin School of Medical Research, Mills Road, Canberra, ACT 2601, Australia

FEATURES
source

Location/Qualifiers
1..852
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/tissue_type="placenta"
1..852
/gene="4SPAN4"
46..798
/gene="4SPAN4"
/codon_start=1
/product="four-span transmembrane protein 4"
/protein_id="AAL56224.1"
/db_xref="GI:18028936"
/translation="MNSMTSAVPVANSVLVAVPHNGVPTPGIMSHVPLYPNSQPVH
LVPNGPPLSVNNGQPVOKALKEKTLGATQIILGLHIGLSGIMATVVLVEYLSIS
FYGGFPWGLWFIISGLSVAENQPSYCLLSGLGLNIYSACISAVGVILFTDL
SIPHPYAPDYVPYAVNGVPMGATISGLLVFCLLEFGIACASSHFGCOLVCCQSSNVS
VIYPNIYAANPVTPEVTPSPSYSSSEIQANK"

BASE COUNT
ORIGIN

195 a 230 c 204 g 223 t
Query Match 6.2%; Score 81.2; DB 9; Length 852;
Best Local Similarity 58.3%; Pred. No. 1e-09;
Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;

QY 590 TCTACTGCTGTTATTTGGTATACCCATCTGGGTGGCCCTTCTTTTATTATCTCTGGC 649
DB 343 TCTATTTTCATCTACGGAGGCTTCCCTCTCTGGGAGGCTTGGTTTATCTATTCAGGA 402
QY 650 TCTCTCTCTGTGTCAGCATCCCAAGAGCT---TTCCCGTGTCTGGTGAAGAGGAGCCTG 706
DB 403 TCTCTCTCCGTGGCAGCAGAAATCAGCCATATCTTATTGCTGCTCTGGCAGTTG 462
QY 707 GGAATGACATGTTAGTTCTATCTTGGCCCTTCATTGGAGTGATTCCTGCTGTTGAT 766
DB 463 GGCTGAACATCGTCAGTGAATCTGCTCGAGTGGAGTCATCTTTCATCACAGAT 522
QY 767 ATGTGCAT-----CAATGGGTAGCTGGCCCAAGACTACTGGCCGCTGTT 811
DB 523 CTAAGTATTCACCACCATATGCTTACCCGACATATCTTACGCTGGGTGTGAAC 582
QY 812 TCTGAAAGGCAATTCAGCCACGCTGATGATCTTCTCCCTCTTGGAGTTCTTCTGAGCT 871
DB 583 CTGGAATGGCGAATTCCTGGCGTGTCTGCTGCTTCTGCTCTCTGCTGAGTTTGGCATCGCA 642
QY 872 TGTGCCACAGCCATTTTGGCAACA 897
DB 643 TGCACATCTTCCCACTTTGGCTGCCA 668

RESULT 12
LOCUS

BC022895
DEFINITION Homo sapiens, membrane-spanning 4-domains, subfamily A, member 8B, clone MGC:27396 IMAGE:4691480, mRNA, complete cds.
ACCESSION BC022895
KEYWORDS BC022895.1 GI:18605601
SOURCE MGC.
ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 1305)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: ccapbs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

FEATURES
source

1..1305
/organism="Homo sapiens"
/db_xref="LocusID:83661"
/db_xref="taxon:9606"
/clone="MGC:27396 IMAGE:4691480"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
129..881
/codon_start=1
/product="membrane-spanning 4-domains, subfamily A, member 8B"
/protein_id="AAH22895.1"
/db_xref="GI:18605602"
/translation="MNSMTSAVPVANSVLVAVPHNGVPTPGIMSHVPLYPNSQPVH
LVPNGPPLSVNNGQPVOKALKEKTLGATQIILGLHIGLSGIMATVVLVEYLSIS
FYGGFPWGLWFIISGLSVAENQPSYCLLSGLGLNIYSACISAVGVILFTDL
SIPHPYAPDYVPYAVNGVPMGATISGLLVFCLLEFGIACASSHFGCOLVCCQSSNVS
VIYPNIYAANPVTPEVTPSPSYSSSEIQANK"

BASE COUNT
ORIGIN

340 a 341 c 287 g 337 t
Query Match 6.2%; Score 81.2; DB 9; Length 1305;
Best Local Similarity 58.3%; Pred. No. 1e-09;
Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;
QY 590 TCTACTGCTGTTATTTGGTATACCCATCTGGGTGGCCCTTCTTTTATTATCTCTGGC 649
DB 426 TCTATTTTCATCTACGGAGGCTTCCCTCTCTGGGAGGCTTGGTTTATCTATTCAGGA 485
QY 650 TCTCTCTCTGTGTCAGCATCCCAAGAGCT---TTCCCGTGTCTGGTGAAGAGGAGCCTG 706
DB 486 TCTCTCTCCGTGGCAGCAGAAATCAGCCATATCTTATTGCTGCTCTGGCAGTTTG 545

RESULT 15	AX369167	861 bp	DNA	linear	PAT 16-FEB-2002
LOCUS	Sequence 1877 from Patent WO0204514.				
DEFINITION	AX369167				
ACCESSION	AX369167.1	GI:18857177			
VERSION					
KEYWORDS	human.				
SOURCE					
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Wang, T., Watanabe, Y., Henderson, R. A., Johnson, J. C., Retter, M. W., Marnerakis, M., Carter, D., Fanger, G. R., Vedvick, T. S., Bangur, C. S., McNabb, A., Fanger, N., Switzer, A., McNell, P. D. and Clapper, J. D. Compositions and methods for the therapy and diagnosis of lung cancer				
JOURNAL	Patent: WO 0204514-A 1877 17-JAN-2002;				
FEATURES	CORINA CORPORATION (US) Location/Qualifiers 1. 861 /organism="Homo sapiens" /db_xref="taxon:9606"				
BASE COUNT	186 a 244 c 209 g 222 t				
ORIGIN					
Query Match	6.0%;	Score 79.6;	DB 6;	Length 861;	
Best Local Similarity	58.0%;	Pred. No. 2.6e-09;			
Matches	189;	Conservative 0;	Mismatches 119;	Indels 18;	Gaps 2;
QY	590	TCTACTGCTGTATTTCGGTATACCATTCCTGGGGTGGCGCTTCTTTTATATCTCTGGC	649		
Db	406	TCTATTTCATTCACGGAGGCTTTCCTCTCTGGGGAGGCTTGTGGTTATCATTTTCAGAA	465		
QY	650	TCTCTCTCTGTGTCACATCCACGAGCT---TTCCCGTTGCTCTGTGAAAGGCAGCCTG	706		
Db	466	TCTCTCTCGTGGCAGCAGAAATACACCATATTCATTTCCTGCTGCTGGCAGTTTG	525		
QY	707	GGAATGAACATTGTTAGTTCTATCTTTGGCCTTCATTGGAGTGAATTCGCTGCTGGGAT	766		
Db	526	GGCTTCAACATCTCAGTGCATCTGCTCTGCAAGTTGGAGTCACTCTTCATCAGACAT	585		
QY	767	ATGTGCAT-----CAATGGGTAGCTGGCCACAGACTACTTGGGCGCTTC	811		
Db	586	CTAAGTATTTCCCCACCCATATGCTTACCCCGACTATTATCTTACGCTGGGTGTGAAC	645		
QY	812	TCTGGAAAAGGCATTTTCAGCCAGCTGATGATCTTCTCCTCTTGGAGTTCCTTCGAGCT	871		
Db	646	CCTGGAATGCGGATTTCTGGCGTGTGCTGGTCTCTCTGCTCTCTGAGTTGGCATCGCA	705		
QY	872	TGTGCCACAGCCCATTTTGGCAACCA	897		
Db	706	TGGCATCTTCCACATTTGGCTGCCA	731		

Search completed: July 16, 2003, 13:27:52
Job time : 2399.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 03:41:41 ; Search time 224.967 Seconds
(without alignments)

13183.634 Million cell updates/sec

Title: US-09-981-353-81

Perfect score: 1317

Sequence: 1 aeggaacaagtagtccat.....ttttctcttaataaatagtc 1317

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1317	100.0	1317	20	AAZ06368 Human receptor pro
2	1108	84.1	1141	21	AAZ95381 Human colon specifi
3	1084.4	82.3	1291	21	AAA26287 Human secreted pro
4	1079	81.9	1292	21	AAA26331 Human secreted pro
5	712.8	54.1	854	23	AAS81191 DNA encoding novel
6	389.6	29.6	441	21	AAC98717 Human colon cancer
7	303.6	23.1	351	22	AAS37436 Novel human diagno
8	292.6	22.2	342	22	AAZ37580 Novel human diagno
9	291.2	22.1	373	22	AAF66008 Novel human polynu

10	288.6	21.9	383	22	AAZ37423	Novel human diagno
11	276	21.0	370	22	AAF65997	Novel human polynu
12	81.2	6.2	1236	19	AAV48114	Nucleotide sequenc
13	81.2	6.2	1336	21	AAZ51573	Human cell surface
14	81.2	6.2	1336	24	ABK39631	CDNA encoding 1985
15	81.2	6.2	1374	21	AAC99077	Human pancreatic c
16	81.2	6.2	1375	24	ABL90599	Human polynucleoti
17	79.6	6.0	861	24	ABK39776	DNA encoding Ral2S
18	79.6	6.0	1155	24	ABK39775	DNA encoding Ral2-
19	79.6	6.0	1353	24	ABK39774	CDNA encoding lung
20	74.8	5.7	932	22	AD19723	Dendritic cell (DC
21	74.8	5.7	945	19	AAV56668	Human secreted pro
22	74.8	5.7	1001	21	AAZ52580	Human secreted pro
23	74.8	5.7	1036	22	AD19722	Dendritic cell (DC
24	74.8	5.7	1106	22	ABA09233	Human secreted pro
25	74.8	5.7	1106	22	AAH99750	Human protein enco
26	74.8	5.7	1108	22	AAF63724	Human HAIERbs-iso
27	74.8	5.7	1289	22	AAH9291	Human cell surface
28	74.8	5.7	1488	20	AZ33586	Human breast tumou
29	74.8	5.7	1522	20	AAZ06369	Human receptor pro
30	74.8	5.7	1570	21	AAA26426	Human secreted pro
31	74.8	5.7	1728	24	ABK34885	Human cDNA encodin
32	74.8	5.7	1743	21	AAF18295	Lung cancer associ
33	74.8	5.7	1871	22	AAH35024	Human colon cancer
34	74.8	5.7	1871	24	ABL90750	Human polynucleoti
35	74.8	5.7	2192	21	AAA27130	Human inflammation
36	74.4	5.6	648	24	ABQ58726	Human colon cancer
37	72	5.5	1060	20	AAZ97730	Extended human sec
38	71.2	5.4	2545	11	AAQ04645	Encodes beta subun
39	71.2	5.4	2545	12	AAQ14734	Encodes beta subun
40	71.2	5.4	2545	14	AAQ51021	Human FcER1 beta g
41	69.2	5.3	708	22	AAF77694	Murine wild-type F
42	61.2	4.6	999	22	AAH99112	Human EST-derived
43	61.2	4.6	1025	23	AAS76894	DNA encoding novel
44	60	4.6	60	24	ABN38453	Human spliced tran
45	60	4.6	489	24	ABL38455	Human colon tumour

ALIGNMENTS

RESULT 1
AAZ06368
ID AAZ06368 standard; DNA; 1317 BP.
XX AC AAZ06368;
XX DT 26-OCT-1999 (first entry)
XX DE Human receptor protein (HURP) 3 nucleotide sequence.
XX KW receptor; cancer; autoimmune disorder; inflammation;
KW antagonist; cell surface protein; cell signalling;
KW antibody; ds.
XX OS Homo sapiens.
XX FH Key
XX CDS 221..1024
FT /*tag= a
FT /product= Human receptor protein 3
XX PN WO9941375-A2.
XX PD 19-AUG-1999.
XX PF 05-FEB-1999; 99WO-US02572.
XX PR 12-FEB-1998; 98US-0022939.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KJ;

PI Hillman JL, Lal P, Shah P, Tang YT, Yue H;
 DR WPI: 1999-494536/41.
 DR P-PSDB; RAY15224.
 XX
 PT New human receptor proteins, used e.g. to treat, prevent and
 PT diagnose gastrointestinal and developmental disorders - and related
 PT nucleic acids, vectors, transformed cells, antibodies, agonists and
 PT antagonists
 XX
 PS Claim 11; Page 88; 94pp; English.
 XX
 CC This is the consensus sequence of the Human receptor protein 3 (HURP-3)
 CC derived from the following overlapping and/or extended nucleic acid
 CC sequences: Incyte clones 611082 (COLN0701), 776770 (COLN0705), 904152
 CC (COLN0707), 908496 (COLN0709), and 1866665 (SKIN0701).
 CC HURP-3 is expressed in cancerous, inflamed, fetal/proliferating, and
 CC reproductive tissue. HURP-3 therefore appears to have a role in some
 CC forms of cancer, autoimmune/inflammatory disorders, reproductive
 CC disorders, and developmental disorders.
 CC This gives rise to the possibility of using an antagonist or an antibody
 CC of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.
 XX
 SQ Sequence 1317 BP; 375 A; 280 C; 278 G; 384 T; 0 other;

Query Match 100.0%; Score 1317; DB 20; Length 1317;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGGAACAAGTAAGTCATGATGACGTTCTTGCTATCTCTCCCAATCAATGGCA 60
 Db 1 AAGGAACAAGTAAGTCATGATGACGTTCTTGCTATCTCTCCCAATCAATGGCA 60
 Qy 61 CAACACTGGTGGTGTCTACCTGTCTGGTCTCTCTAGATGGAGGATGAAGCAA 120
 Db 61 CAACACTGGTGGTGTCTACCTGTCTGGTCTCTCTAGATGGAGGATGAAGCAA 120
 Qy 121 GTTCTTGACTATGTTGAGGCCAGTTGAAAATGAGGGAGATTAACCAATGAACGAAA 180
 Db 121 GTTCTTGACTATGTTGAGGCCAGTTGAAAATGAGGGAGATTAACCAATGAACGAAA 180
 Qy 181 CAAGAAGAACAACAAGAGAGAGATGAAGAGATATGATGATCCAGCCCAAC 240
 Db 181 CAAGAAGAACAACAAGAGAGAGATGAAGAGATATGATGATCCAGCCCAAC 240
 Qy 241 AAGCCATGCTGAAGTAAATGAACCATACCAACCCCTTACCCACAGCAGCTTTATGGC 300
 Db 241 AAGCCATGCTGAAGTAAATGAACCATACCAACCCCTTACCCACAGCAGCTTTATGGC 300
 Qy 301 TCCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGTTGCTCA 360
 Db 301 TCCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGTTGCTCA 360
 Qy 361 GCGTGGCTCAGCCCTACGGCATCATCTCCGGGATCTTCTAGCAGTCAACCGGTCA 420
 Db 361 GCGTGGCTCAGCCCTACGGCATCATCTCCGGGATCTTCTAGCAGTCAACCGGTCA 420
 Qy 421 AGGAATATACAAATGATAATCCAAAGTGGGACACAGCAGTAATGAACCTTAAAGAGA 480
 Db 421 AGGAATATACAAATGATAATCCAAAGTGGGACACAGCAGTAATGAACCTTAAAGAGA 480
 Qy 481 AGCAAAAGCCTAGGGGTGATCCAGATCATGTTGGATTGATGCACATTTGTTGGAAAT 540
 Db 481 AGCAAAAGCCTAGGGGTGATCCAGATCATGTTGGATTGATGCACATTTGTTGGAAAT 540
 Qy 541 TGTGTTGTTTAAATATCTCTCTTTTAGAGAGATATTAGGTTTGCCCTACTGCTGT 600
 Db 541 TGTGTTGTTTAAATATCTCTCTTTTAGAGAGATATTAGGTTTGCCCTACTGCTGT 600
 Qy 601 TATTGGTGATACCCATCTCGGGTGGCTTTCTTTTATTATCTCTGGCTCTCTCTGT 660
 Db 601 TATTGGTGATACCCATCTCGGGTGGCTTTCTTTTATTATCTCTGGCTCTCTCTGT 660

Qy 661 GTACGATCCCAAGGAGCTTTCCCGTGTGTCTGGTGAAGGAGCCTGGGAATGAACATGT 720
 Db 661 GTACGATCCCAAGGAGCTTTCCCGTGTGTCTGGTGAAGGAGCCTGGGAATGAACATGT 720
 Qy 721 TAGTTCATCTTGGCCCTTATGAGTGATTTCTGCTGTGGATATGTCATCAATGG 780
 Db 721 TAGTTCATCTTGGCCCTTATGAGTGATTTCTGCTGTGGATATGTCATCAATGG 780
 Qy 781 GGTAGCTGGCCAAGCACTACTGGGCGCTCTTCTGGAAAAGGCAATTTTCAGGCACGCTGAT 840
 Db 781 GGTAGCTGGCCAAGCACTACTGGGCGCTCTTCTGGAAAAGGCAATTTTCAGGCACGCTGAT 840
 Qy 841 GATCTTCTCCCTCTTGGAGTTCCTTCTGAGTCTTGTGCCACACGCCCATTTTGGCAACCAAGC 900
 Db 841 GATCTTCTCCCTCTTGGAGTTCCTTCTGAGTCTTGTGCCACACGCCCATTTTGGCAACCAAGC 900
 Qy 901 AAACACCAACCAATATGCTGCTGCTGGTTATTCCCAATATGATGAAGCAACCCCTGT 960
 Db 901 AAACACCAACCAATATGCTGCTGCTGGTTATTCCCAATATGATGAAGCAACCCCTGT 960
 Qy 961 GACACCAAGCTCTTCTCAGCTCCTCCAGATGCAACAACCTACTCAGCTAATGCCCTTAA 1020
 Db 961 GACACCAAGCTCTTCTCAGCTCCTCCAGATGCAACAACCTACTCAGCTAATGCCCTTAA 1020
 Qy 1021 ATAGTAAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACCTACTTGCAAAACCT 1080
 Db 1021 ATAGTAAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACCTACTTGCAAAACCT 1080
 Qy 1081 TCCTTAAGAGATGCTCTTTATTGTTCTACAATGATTTCTAGTCTTTAAAACCTGTGTTGA 1140
 Db 1081 TCCTTAAGAGATGCTCTTTATTGTTCTACAATGATTTCTAGTCTTTAAAACCTGTGTTGA 1140
 Qy 1141 GATTTGTTTTAGGTTGGTCTGCTAATGATGGCTGTATCTCCCTTCACTGTCTCTTCTCTAC 1200
 Db 1141 GATTTGTTTTAGGTTGGTCTGCTAATGATGGCTGTATCTCCCTTCACTGTCTCTTCTCTAC 1200
 Qy 1201 ATTACCACTACTACATGCTGGCAAGGTGAAGAGTACAGAGACTGAAAAATGATTTCTGCA 1260
 Db 1201 ATTACCACTACTACATGCTGGCAAGGTGAAGAGTACAGAGACTGAAAAATGATTTCTGCA 1260
 Qy 1261 ACTCTCTTAAAGTTAGAAATGTTCTTCTCATATTTACTTTTCTCTTAAATAAATGTC 1317
 Db 1261 ACTCTCTTAAAGTTAGAAATGTTCTTCTCATATTTACTTTTCTCTTAAATAAATGTC 1317

RESULT 2
 AAZ95381
 ID AAZ95381 standard; DNA; 1141 BP.
 XX
 AC AAZ95381;
 XX
 DT 01-JUN-2000 (first entry)
 XX
 DE Human colon specific gene (CSG) SEQ ID NO:3.
 XX
 KW Human; colon specific gene; CSG; diagnosis; colon cancer; detection;
 XX cytosolic; ds.
 OS Homo sapiens.
 XX
 PN WO200007632-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF 20-JUL-1999; 99WO-US16357.
 XX
 PR 04-AUG-1998; 98US-0095231.
 XX
 PA (DIAD-) DIADEXUS LLC.
 XX
 PI Sun Y, Recipon H, Macina RA;
 XX
 DR WPI; 2000-205579/18.

XX Novel methods for diagnosing, monitoring, staging, imaging and treating
PT colon cancer by measuring the level of colon specific gene markers
PS
PS
XX Claim 6; Page 38-39; 42pp; English.

CC The present invention describes a method for diagnosing the presence of
CC colon cancer in a patient. The method comprises measuring levels of
CC colon specific gene markers (CSG) in cells, tissues or bodily fluids,
CC and comparing the measured levels of CSG with levels of CSG from a
CC normal human control, where an increase in measured CSG levels in the
CC patient versus control is associated with the presence of colon cancer.
CC AA295379 to AA295381 represent human CSG sequences. The method is used
CC to detect, monitor, stage or give a prognosis for colon cancer.
CC Antibodies against CSGs are used for detection or image localisation of
CC the CSGs. The antibodies can be conjugated to cytotoxic agent or drug and
CC used to treat colon cancer. The method of the invention is more accurate
CC than prior art clinical methods for staging colon cancer, because it
CC measures colon specific markers, and, unlike pathological staging
CC methods, do not depend on an invasive procedure.

XX
SQ Sequence 1141 BP; 317 A; 248 C; 238 G; 338 T; 0 other;

Query Match 84.1%; Score 1108; DB 21; Length 1141;
Best Local Similarity 98.2%; Pred. No. 2.4e-285;
Matches 1120; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 178 AAACAAGAGAAACAAACAGAGAGGAAATGAAAAGACATAAATGATCTATCCAGCC 237
DB 2 AGAGAAGAGGAAACATAGAGGTGCCAAGGACAAAGACATAAATGATCTATCCAGCC 61

QY 238 AACAGCCATGCTGAAGTAAATGAACACCATACCCACCCCTTACCCACCAAGCAGCTTTAT 297
DB 62 AACAGCCATGCTGAAGTAAATGAACACCATACCCACCCCTTACCCACCAAGCAGCTTTAT 121

QY 298 GGCTCTGGATTTCACAGCCCTCTGGGTTCAATCAACTTAGAAGCAAGCAAGCTCAGGGTGC 357
DB 122 GGCTCTGGATTTCACAGCCCTCTGGGTTCAATCAACTTAGAAGCAAGCAAGCTCAGGGTGC 181

QY 358 TCAGGCTGCTCAGCCCTACGGGATCATCTCCGGGAATCTTGGTAGCAGTCAACCCGG 417
DB 182 TCAGGCTGCTCAGCCCTACGGGATCATCTCCGGGAATCTTGGTAGCAGTCAACCCGG 241

QY 418 TCAAGGAAATATACAAATGATAAATCAAGTGTGGGAACAGCAGTGAATGAACCTTTAAGA 477
DB 242 TCAAGGAAATATACAAATGATAAATCAAGTGTGGGAACAGCAGTGAATGAACCTTTAAGA 301

QY 478 AGAAGCAAGGCACTAGGGGTGATCCAGATCATGTTGGATGATGATCCATTTGGTTGG 537
DB 302 AGAAGCAAGGCACTAGGGGTGATCCAGATCATGTTGGATGATGATCCATTTGGTTGG 361

QY 538 AATTGTTTGTGTTTAAATATCTCTCTCTTTAGAGAAGTATTAGGTTTGGCTCTACTGC 597
DB 362 AATTGTTTGTGTTTAAATATCTCTCTCTTTAGAGAAGTATTAGGTTTGGCTCTACTGC 421

QY 598 TGTATTGTGGATACCCATTCGGGGTGGCCCTTTCTTTTATATCTCTGGCTCTCTCTC 657
DB 422 TGTATTGTGGATACCCATTCGGGGTGGCCCTTTCTTTTATATCTCTGGCTCTCTCTC 481

QY 658 TGTGTCAGCATCCAAGGACCTTCCCGTTGCTGGTGAAGGAGGAGGAGGAGTGAACAT 717
DB 482 TGTGTCAGCATCCAAGGACCTTCCCGTTGCTGGTGAAGGAGGAGGAGTGAACAT 541

QY 718 TGTGTTCTATCTTGGCTTCAATGAGTGTCTCTGCTGGTGGATATGTCATCAA 777
DB 542 TGTGTTCTATCTTGGCTTCAATGAGTGTCTCTGCTGGTGGATATGTCATCAA 601

QY 778 TGGGTAGCTGGCCAGACTACTGGGCCGCTGCTTTCTGGAAAAGGAGGATTTAGCCACGCT 837
DB 602 TGGGTAGCTGGCCAGACTACTGGGCCGCTGCTTTCTGGAAAAGGAGGATTTAGCCACGCT 661

QY 838 GATGATCTCTCCCTCTGGAGTCTTCTGCTAGCTTGTGTCACAGGCCATTTTGGCCAAACA 897
DB 1141

DB 662 GATGATCTCTCCCTCTTGGAGTCTTCTGAGTCTTCTGCCACAGCCATTTTGGCCAAACA 721

QY 898 AGCAAAACACCAACCAATATGTCCTCTCTGGTATTTCACAAATATGATGTAAGAGCAACCC 957

DB 722 AGCAAAACACCAACCAATATGTCCTCTCTGGTATTTCACAAATATGATGTAAGAGCAACCC 781

QY 958 TGTGACACAGCGTCTTCTTCAAGTCTCTCCAGATCCCAACAACTACTAGCTAATGCCCC 1017

DB 782 TGTGACACAGCGTCTTCTTCAAGTCTCTCCAGATCCCAACAACTACTAGCTAATGCCCC 841

QY 1018 TAAATAGTAAAGAAAAAGGGGTATCAGTCTATCTCATCTCATGGAGAAAACTACTTGCANAA 1077

DB 842 TAAATAGTAAAGAAAAAGGGGTATCAGTCTATCTCATCTCATGGAGAAAACTACTTGCANAA 901

QY 1078 ACTTCTTAAAGAGATGCTCTTTTATTCTCTACATGATTTTCTAGTCTTTTAAAAACTGTGTT 1137

DB 902 ACTTCTTAAAGAGATGCTCTTTTATTCTCTACATGATTTTCTAGTCTTTTAAAAACTGTGTT 961

QY 1138 TGAGATTTGTTTTAGGTTGGTGGCTTAATGATGGCTGATCTCTCCCTTCACTGTCTCTTCC 1197

DB 962 TGAGATTTGTTTTAGGTTGGTGGCTTAATGATGGCTGATCTCTCCCTTCACTGTCTCTTCC 1021

QY 1198 TACATTACCACACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAAATGATTCT 1257

DB 1022 TACATTACCACACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAAATGATTCT 1081

QY 1258 GCAACTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCTTCTTAATAAATGTC 1317

DB 1082 GCAACTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCTTCTTAATAAATGTC 1141

RESULT 3
AAA26287
ID AAA26287 standard; cDNA; 1291 BP.
XX
AC AAA26287;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein gene 7 SEQ ID NO:17.
XX
KW Human; secreted protein; diagnosis; neuroprotective; neurotropic;
KW neuroleptic; antinaptic; cerebroprotective; immunomodulatory;
KW anti-microbial; cardiant; cytostatic; antinflammatory; haemostatic;
KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
KW neural; reproductive; immune disorder; immunodeficiency; infection;
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
KW ischaemia; mania; dementia; obsessive compulsive disorder;
KW viral prophylaxis; developmental disorder; sexually-linked disorder;
KW cardiovascular disorder; food additive; preservative; ss.
OS Homo sapiens.
XX
XX WO200011014-A1.
PN
XX
XX
PD 02-MAR-2000.
XX
XX 24-AUG-1999; 99WO-US19330.
PF
XX 25-AUG-1998; 98US-0097917.
PR
PR 31-AUG-1998; 98US-0098634.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G;
PI Duan RD;
XX
XX WPI; 2000-224656/19.
DR P-PSDB; AAY91352.
XX
XX

PT Novel secreted proteins and corresponding DNA molecules that can be
PT Alzheimers', cancer, and immune disorders -
XX
PS Claim 1; Page 315-316; 416pp; English.
XX

CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the
CC human secreted proteins given in AAY91346 to AAY91449. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: neuroprotective; neurotropic;
CC neuroleptic; antimalic; cerebroprotective; immunomodulatory;
CC anti-microbic; cardiant; cytostatic; antinflammatory; haemostatic;
CC anticonvulsant; and vasotropic. The polynucleotides and proteins may be
CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
CC or gene therapy. Conditions treatable by the proteins of the invention
CC include neural, reproductive, or immune disorders, especially
CC immunodeficiency, infection, lymphomas, demyelinating diseases,
CC auto-immunities, cancer, general microbial infection, inflammation,
CC aneurysms and haemorrhages. Specific examples include: Alzheimer's
CC disease; Parkinson's; Huntington's; Tourette syndrome; multiple
CC sclerosis; meningitis; ischaemia; prostate cancer; mania; dementia;
CC obsessive compulsive disorder and viral prophylaxis. The polynucleotides
CC and proteins can also be used in the detection of disorders associated
CC with the function of the protein, for example, the detection of
CC developmental disorders, sexually-linked disorders, or disorders of the
CC cardiovascular system. They may also be used as food additives or
CC preservatives. AAA26272 to AAA26280 and AAY91345 are sequences used in
CC the exemplification of the present invention.
XX
XX Sequence 1291 BP; 371 A; 285 C; 272 G; 358 T; 5 other;

Query Match 82.3%; Score 1084.4; DB 21; Length 1291;
Best Local Similarity 97.8%; Pred. No. 5.1e-279;
Matches 1111; Conservative 1; Mismatches 17; Indels 7; Gaps 1;

QY 182 AGCAATGCTGAAGTAAATGAACCATACCCACCTTACCCACCAAGCAGCTTTATGGCT 301
DB 75 AAGAGGAACATAGAGTGCCAAAGGAACAAAGACATAATGATGTCATCCAAAGCAACA 134
QY 242 AGCCATGCTGAAGTAAATGAACCATACCCACCTTACCCACCAAGCAGCTTTATGGCT 301
DB 135 AGCCATGCTGAAGTAAATGAACCATACCCACCTTACCCACCAAGCAGCTTTATGGCT 194
QY 302 CTGTGATTTCAACAGCCTCTGGTTCAATCAATCAATCAATCAATCAATCAATCAAT 361
DB 195 CTGTGATTTCAACAGCCTCTGGTTCAATCAATCAATCAATCAATCAATCAATCAAT 254
QY 362 CGTGTCTAGCCCTACGGGATCACATCTCCGGGATCTTTGCTAGCAGTCAACCGGTC 421
DB 255 CGTGTCTAGCCCTACGGGATCACATCTCCGGGATCTTTGCTAGCAGTCAACCGGTC 314
QY 422 GGAATATACAAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTAAAGAGAA 481
DB 315 GGAATATACAAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTAAAGAGAA 374
QY 482 GCAAGGACCTAGGGGATCCAGATCATGTTGGATTGATGACATTTGGTTTGGAAAT 541
DB 375 GCAAGGACCTAGGGGATCCAGATCATGTTGGATTGATGACATTTGGTTTGGAAAT 434
QY 542 GTTTTGTGTTAAATCTCTCTTTTAGAAGATTAATAGGTTTGGCTCTACTGCTTT 601
DB 435 GTTTTGTGTTAAATCTCTCTTTTAGAAGATTAATAGGTTTGGCTCTACTGCTTT 494
QY 602 ATTTGTGATACCCATCTGGGGTGGCTTTCTTTTATTTATCTCTGCTCTCTCTCTG 661
DB 495 ATTTGTGATACCCATCTGGGGTGGCTTTCTTTTATTTATCTCTGCTCTCTCTCTG 554
QY 662 TCAGCATCAAGGAGCTTTCCCGTTGCTGTGTGAAGGACGCTGGGAATGAACATTTGT 721
DB 555 TCAGCATCAAGGAGCTTTCCCGTTGCTGTGTGAAGGACGCTGGGAATGAACATTTGT 614
QY 722 AGTTCTATCTTGGCTTTCATTTGGAGTGATTTCTGCTGCTGGTGATGATGTCATCAATGG 781

DB 615 AGTTCTATCTTGGGCTTCATTTGGAGTGATTTGCTGCTGGTGATATGTGTCATCAATGGG 674
QY 782 GTAGTGTGCCAAGACTACTGCGGCGCTGCTTTCTGGAAGAGCATTTTACACCGCAGCTGATG 841
DB 675 GTAGTGTGCCAAGACTACTGCGGCGCTGCTTTCTGGAAGAGCATTTTACACCGCAGCTGATG 734
QY 842 ATCTTCTCCCTCTTGGAGTCTTCTGAGCTTGTGCCACAGCCCATTTTGGCCAAACAACA 901
DB 735 ATCTTCTCCCTCTTGGAGTCTTCTGAGCTTGTGCCACAGCCCATTTTGGCCAAACAACA 794
QY 902 AACACCAACAACAATATCTGCTGCTGCTTATTCACAAATATGATGAAGCAACCCCTG 961
DB 795 AACACCAACAACAATATCTGCTGCTGCTTATTCACAAATATGATGAAGCAACCCCTG 854
QY 962 ACACAGAGCTCTTCTTTCAGCTCCCTCCAGATGCAACAACACTACTCAGCTTAATGCCCTAAA 1021
DB 855 ACACAGAGCTCTTCTTTCAGCTCCCTCCAGATGCAACAACACTACTCAGCTTAATGCCCTAAA 910
QY 1022 TAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGAGAGAAAACACTACTTGCAGAACTT 1081
DB 911 --TAAAGAAAAGGGGTATCAGTCTAATCTCATGAGAGAAAACACTACTTGCAGAACTT 967
QY 1082 CTTAAGAGAGTCTCTTTTATTTGCTTACAAATGATTTCTAGTCTTTTAAACACTGCTGTCAG 1141
DB 968 CTTAAGAGAGTCTCTTTTATTTGCTTACAAATGATTTCTAGTCTTTTAAACACTGCTGTCAG 1027
QY 1142 ATTTGTTTTAGTTGGTGGCTTAATGATGAGTGTATCTCCCTTCACTGCTCTCTCTTCTACA 1201
DB 1028 ATTTGTTTTAGTTGGTGGCTTAATGATGAGTGTATCTCCCTTCACTGCTCTCTCTTCTACA 1087
QY 1202 TTACACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAATGATTTCTGCA 1261
DB 1088 TTACACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAATGATTTCTGCA 1147
QY 1262 CTCCTCTAAGTTAGAAATGTTCTGTTTCATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1317
DB 1148 CTCCTCTAAGTTAGAAATGTTCTGTTTCATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1203

RESULT 4
AAA26331
ID AAA26331 standard; cDNA; 1292 BP.
XX AAA26331;
XX
XX 29-JUN-2000 (first entry)
XX
XX Human secreted protein gene 7 SEQ ID NO:61.
DE
XX
KW Human; secreted protein; diagnosis; neuroprotective; neurotropic;
KW neuroleptic; antimalic; cerebroprotective; immunomodulatory;
KW anti-microbic; cardiant; cytostatic; antinflammatory; haemostatic;
KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
KW neural; reproductive; immune disorder; immunodeficiency; infection;
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
KW ischaemia; mania; dementia; obsessive compulsive disorder;
KW viral prophylaxis; developmental disorder; sexually-linked disorder;
KW cardiovascular disorder; food additive; preservative; ss.
XX
XX Homo sapiens.
OS
XX WO200011014-A1.
PN
XX
XX
PD 02-MAR-2000.
XX
XX 24-AUG-1999; 98WO-US19330.
PF
XX 25-AUG-1998; 98US-0097917.
PR
XX 31-AUG-1998; 98US-0098634.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
DR WPI; 2001-530177/58.
XX
XX New polynucleotides and polypeptides, useful for diagnosis and
PT treatment of breast, lung and colon cancer -
XX
XX Claim 1; Page 756; 1193pp; English.
XX
XX The invention relates to new polynucleotides and polypeptides, useful for
CC diagnosis and treatment of breast, lung and colon cancer. The sequences
CC can be used in detecting differentially expressed genes correlated with a
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample derived from a
CC cell suspected of being cancerous. They can also be used to inhibit
CC tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
CC sequences of the invention.
XX
XX Sequence 342 BP; 119 A; 78 C; 78 G; 65 T; 2 other;

Query Match 22.2%; Score 292.6; DB 22; Length 342;
Best Local Similarity 92.5%; Pred. No. 5.1e-68;
Matches 307; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 172 TGAAGCAACAGGAAAGAAACAAACAGAGAGGAGTAAAGAGACATATGATGTCATC 231
Db TGGAGCAGAGAGAGGAGGAAACATAGAGGTGCCAAGGAGAAAGACATATGATGTCATC 69
QY 232 CAAGCCAAACAGCCATGCTGAAGTAAATGAACCATACCCCAACCTTACCCCAAGCAG 291
Db TGGAGCAGAGAGAGGAGGAAACATAGAGGTGCCAAGGAGAAAGACATATGATGTCATC 129
QY 292 CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCAATCACTTAGAAAACCAAGCTCA 351
Db CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCAATCACTTAGAAAACCAAGCTCA 189
QY 352 GGGTGTCTCAGCGTCTCAGCCCTAGCGCATCACATCTCCGGGAATCTTTGTAGCAGTCA 411
Db GGGTGTCTCAGCGTCTCAGCCCTATGGCATCACATCTCCGGGAATCTTTGTAGCAGTCA 249
QY 412 ACCGGGTCAAGGAAATATACAAATGATAATCCAAAGTGTGGGACAGCAGTAATGAACCTT 471
Db ACCGGGTCAAGGAAATATATTAATGATAATCCAAAGTGTGGGACAGCAGTAATGAACCTT 309
QY 472 TAAAGAGAAGCAAGGCACTAGGGGTGATCC 503
Db TAAAGAGAAGCAAGGCACTAGGGGTGATCC 341

RESULT 9
AAF66008
ID AAF66008 standard; cDNA: 373 BP.
XX
AC AAF66008;
XX
DT 09-APR-2001 (first entry)
DE
XX Novel human polynucleotide, SEQ ID NO: 1764.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX

30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkvenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 796; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
XX Sequence 373 BP; 127 A; 84 C; 90 G; 72 T; 0 other;
Query Match 22.1%; Score 291.2; DB 22; Length 373;
Best Local Similarity 91.7%; Pred. No. 1.2e-67;
Matches 308; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 172 TGAAGCAACAGGAAAGAAACAAACAGAGAGGAGTAAAGAGACATATGATGTCATC 231
Db TGGAGCAGAGAGAGGAGGAAATCATAGAGGTGCCAAGGAGAAACAGACATATGATGTCATC 96
QY 232 CAAGCCAAACAGCCATGCTGAAGTAAATGAACCATACCCCAACCTTACCCCAAGCAG 291
Db CAAGCCAAACAGCCATGCTGAAGTAAATGAACCATACCCCAACCTTACCCCAAGCAG 156
QY 292 CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCAATCACTTAGAAAACCAAGCTCA 351
Db CTTTATGGCTCCTGGATTTCAACAGCCCTCTGGTTCAATCACTTAGAAAACCAAGCTCA 216
QY 352 GGGTGTCTCAGCGTCTCAGCCCTAGCGCATCACATCTCCGGGAATCTTTGTAGCAGTCA 411
Db GGGTGTCTCAGCGTCTCAGCCCTATGGCATCACATCTCCGGGAATCTTTGTAGCAGTCA 276
QY 412 ACCGGGTCAAGGAAATATACAAATGATAATCCAAAGTGTGGGACAGCAGTAATGAACCTT 471
Db ACCGGGTCAAGGAAATATATTAATGATAATCCAAAGTGTGGGACAGCAGTAATGAACCTT 336
QY 472 TAAAGAGAAGCAAGGCACTAGGGGTGATCCAGAT 507
Db TAAAGAGAAGCAAGGCACTAGGGGTGATCCAGAT 372
RESULT 10
AAS37423
ID AAS37423 standard; cDNA: 383 BP.
XX
AC AAS37423;

QY 352 GGGTCTCAGCGTCTCAGCCCTACGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCA 411
 |||||
 Db 217 GGGTCTCAGCGTCTCAGCCCTATGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCA 276
 |||||
 QY 412 ACCGGGTCAGGAATATACAAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACCTT 471
 |||||
 Db 277 ACCGGGCAAGGAATATATAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACCTT 336
 |||||
 QY 472 TAAAGAAGCAAGCAAG 487
 |||||
 Db 337 TAAAGAAGCAAGCAAG 352
 |||||

RESULT 12

AAV48114

ID AAV48114 standard; cDNA: 1236 BP.

XX

AC AAV48114;

XX

DT 27-OCT-1998 (first entry)

XX

DE Nucleotide sequence encoding clone HTPF86 of TM4SF superfamily.

XX

KW ds; human; receptor; immune disorder; cancers; blood disorder;

KW juvenile rheumatoid arthritis; Graves disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 73..825

FT

FT /*tag= a

FT /product= "Clone HTPF86"

XX

PN W09831799-A2.

XX

PD 23-JUL-1998.

XX

PF 21-JAN-1998; 98WO-US00959.

XX

PR 21-JAN-1997; 97US-0034205.

XX

PR 21-JAN-1997; 97US-0034204.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Gentz RL, Ni J, Rosen CA;

XX

DR WPI; 1998-427559/36.

XX

DR P-PSDB; AAW61619.

XX

PT New isolated poly:nucleotide(s) and encoded receptor poly:peptide(s)

PT - used to develop products for diagnosing or treating e.g. immune

PT disorders, cancers, blood disorders or immuno-compromised disease

PT states

XX

PS Claim 2; Page 38-39; 79pp; English.

XX

CC Clone HTPF86 is a member of the TM4SF receptor superfamily. The
 CC products generated using the receptor can be used for treating abnormal
 CC conditions related to both an excess of and insufficient amounts of
 CC receptor activity. They can be used in the treatment of e.g. immune
 CC disorders, cancers, blood disorders, juvenile rheumatoid arthritis,
 CC Graves disease or immunocompromised disease states. The products can
 CC also be used for detection and diagnosis.

SQ Sequence 1236 BP; 315 A; 320 C; 276 G; 325 T; 0 other;

Query Match 6.2%; Score 81.2; DB 19; Length 1236;

Best Local Similarity 58.3%; Pred. No. 2.3e-11;

Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;

QY 590 TCTACTGCTGTATTGGTGGATACCCATTCGGGGTGCCTTCTTTATTATCTCTGGC 649

|||||

Db 370 TCTATTTCATTTCTACGGAGGCTTCCCTTCTGGGAGGCTTGTGGTTATCATTTACGA 429
 QY 650 TCTCTCTCTGTGTGTCAGCATCAAGAGCT---TTCCCCGTTTCTGGTGAAGGCGAGCCTG 706
 |||||
 Db 430 TCTCTCTCCGTCGGCAGCAGAAAATCAGCCATATCTTATTCCTGCTGCTGGCAGTTTG 489
 |||||
 QY 707 GGAATGAACATTTAGTTCTATCTTGGCCCTTCATTTGGAGTGATTTCTGCTGCTGGTGGAT 766
 |||||
 Db 490 GCCTTGAACATCGTCAGTGCAATCTGCTCTCAGTTGGAGTTCATCTTTCATCACAGAT 549
 |||||
 QY 767 ATGTGCAT-----CAATGGGGTAGCTGCCCAAGACTACTGGCGCTGCTT 811
 |||||
 Db 550 CTAAGTATTCCTCCACCCCATATGCTTACCCCACTATTATCTTACGCTGGGTGTGAAC 609
 |||||
 QY 812 TCTGAAAAAGGCATTTTCAGCCACGCTGATGATCTTCTCCCTCTTGGAGTTCTTCTAGCT 871
 |||||
 Db 610 CCTGGAATGGCATTTCTGGCGTCTGCTGCTTCTTCTGCTTCTGAGTTTGGCATCGCA 669
 |||||
 QY 872 TGTGCCACAGCCCATTTTGGCAACCA 897
 |||||
 Db 670 TGGCATCTTCCCACTTTGGCTGCCA 695
 |||||

RESULT 13

AAZ51573

ID AAZ51573 standard; cDNA: 1336 BP.

XX

AC AAZ51573;

XX

DT 21-JUN-2000 (first entry)

XX

DE Human cell surface immunomodulator-2 (CSIMM-2) encoding cDNA.

XX

KW Cell surface immunomodulator-2; CSIMM-2; cell proliferation; CD20;
 KW differentiation; signal transduction; drug screening; prevention;
 KW treatment; cancer; leukaemia; melanoma; immune disorder; AIDS;
 KW rheumatoid arthritis; asthma; atherosclerosis; diabetes mellitus;
 KW emphysema; irritable bowel syndrome; multiple sclerosis; diagnosis;
 KW osteoporosis; psoriasis; microbial infections; cytostatic; antiarthritic;
 KW antiasthmatic; immunosuppressive; antiarteriosclerotic; anti-HIV;
 KW antidiabetic; antiinflammatory; neuroprotective; osteopathic;
 KW antipsoriatic; antimicrobial; human; ss.

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 174..926

FT

FT /*tag= a

FT /product= "Cell surface immunomodulator-2"

FT

FT sig_peptide 174..239

FT

FT /*tag= b

FT mat_peptide 240..923

FT

FT /*tag= c

FT

FT /product= "Mature cell surface immunomodulator-2"

FT

FT misc_binding 801..830

FT

FT /*tag= d

FT

FT /bound_moiety= "Hybridisation probe"

XX

XX WO200011150-A1.

PN

XX 02-MAR-2000.

PD

XX 24-AUG-1999; 99WO-US19386.

PF

XX 25-AUG-1998; 98US-0139687.

PR

XX (INCY-) INCYTE PHARM INC.

PA

XX Lal P, Corley NC, Gorgone GA, Guegler KJ, Patterson C, Baughn MR;

PI

XX WPI; 2000-246561/21.

DR

XX P-PSDB; AAY70432.

PT New human cell surface immunomodulatory polypeptides and
PT polynucleotides useful for diagnosis, prevention and treatment of
PT cancer and immune disorders
XX
PS
PS Claim 9; Page 63-64; 70pp; English.
XX
XX The present sequence is a cDNA identified in Incyte clone 2721336,
CC derived from lung tumour cDNA library (LUNGUT10). It encodes
CC human cell surface immunomodulator-2 (CSIMM-2), which is a
CC regulator of cell proliferation, differentiation, cell-cell
CC communication and signal transduction. It shows homology to cell surface
CC antigen, CD20. CSIMM can be used for drug screening, prevention and
CC treatment of cancers such as leukaemia and melanoma, immune disorders
CC such as AIDS, rheumatoid arthritis, asthma, atherosclerosis, diabetes
CC mellitus, emphysema, irritable bowel syndrome, multiple sclerosis,
CC osteoporosis, psoriasis and microbial infections. CSIMM polynucleotide
CC may be used for diagnosis of CSIMM-associated diseases and as source of
CC probes useful in mapping naturally occurring genomic sequences.
XX
XX Sequence 1336 BP; 336 A; 354 C; 296 G; 350 T; 0 other;

Query Match 6.2%; Score 81.2; DB 21; Length 1336;
Best Local Similarity 58.3%; Pred. No. 2.3e-11;
Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;
QY 590 TCTACTGCTGTTATTTGGTGATACCCATTTCTGGGTGGCTTTCTTTTATTATCTCTGGC 649
DB 471 TCTATTTCATTTCTACGGAGGCTTCCCTTCTGGGAGGCTTGTGTTTATCATCTTCAGGA 530
QY 650 TCTCTCTCTGTCTAGCATCCCAAGGAGCT---TTCCCGTTGTCTGTGTAAGGACGCTTG 706
DB 531 TCTCTCTCGTGGGACAGAAAATCAGCCATATCTTATTTGCTGTCTGGCAGTTTG 590
QY 707 GGAATGAACATTTGTTAGTTCTATCTTGGCTTCTATTTGGAGTGATCTGCTGTGGTGGAT 766
DB 591 GGCTTGAACATCTCAGTGCAATCTGCTGAGTTGGAGTCTACTTTCATCATCAGAT 650
QY 767 ATGTGCAT-----CAATGGGGTAGTGGCCAAAGACTACTTGGCCCGTGTCT 811
DB 651 CTAAGTATTTCCTCCACCCCATATGCTTACCCGAGCTATTATCTTACGCTGGGGTGTGAAC 710
QY 812 TCTGGAAGGAGCATTTTCAGCCAGCTGATGATCTTCTCCCTCTTGGAGTTCTTCGTAGCT 871
DB 711 CCTGGAATGGCGATTCTGGCGTGTCTGTCTTCTGCTCTCTGGAGTTTGGCATCGCA 770
QY 872 TGTGCCACAGCCCATTTTCCCAACCA 897
DB 771 TCGGCATCTTCCCACTTTGGCTGCCA 796

RESULT 14

ABK39631
ID ABK39631 standard; cDNA; 1336 BP.

XX ABK39631;

XX 21-MAY-2002 (first entry)

XX cDNA encoding L985P lung tumour associated protein.

XX Lung tumour; cancer; T cell; Immune response stimulator;
KW cytostatic; gene; ss.

XX Homo sapiens.

XX WO200204514-A2.

XX 17-JAN-2002.

XX 10-JUL-2001; 2001WO-US22058.

XX 11-JUL-2000; 2000US-0614124.

XX 29-AUG-2000; 2000US-0651563.

PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI Mcnabb A, Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX WPI; 2002-164634/21.
DR P-PSDB; AAU85539.
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein
XX
XX Claim 7; SEQ ID No 1676; 223pp; English.
XX The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1336 BP; 336 A; 354 C; 296 G; 350 T; 0 other;

Query Match 6.2%; Score 81.2; DB 24; Length 1336;
Best Local Similarity 58.3%; Pred. No. 2.3e-11;
Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;
QY 590 TCTACTGCTGTTATTTGGTGATACCCATTTCTGGGTGGCTTTCTTTTATTATCTCTGGC 649
DB 471 TCTATTTCATTTCTACGGAGGCTTCCCTTCTGGGAGGCTTGTGTTTATCATCTTCAGGA 530
QY 650 TCTCTCTCTGTCTAGCATCCCAAGGAGCT---TTCCCGTTGTCTGTGTAAGGACGCTTG 706
DB 531 TCTCTCTCGTGGGACAGAAAATCAGCCATATCTTATTTGCTGTCTGGCAGTTTG 590
QY 707 GGAATGAACATTTGTTAGTTCTATCTTGGCTTCTATTTGGAGTGATCTGCTGTGGTGGAT 766
DB 591 GGCTTGAACATCTCAGTGCAATCTGCTGAGTTGGAGTCTACTTTCATCATCAGAT 650
QY 767 ATGTGCAT-----CAATGGGGTAGTGGCCAAAGACTACTTGGCCCGTGTCT 811
DB 651 CTAAGTATTTCCTCCACCCCATATGCTTACCCGAGCTATTATCTTACGCTGGGGTGTGAAC 710
QY 812 TCTGGAAGGAGCATTTTCAGCCAGCTGATGATCTTCTCCCTCTTGGAGTTCTTCGTAGCT 871
DB 711 CCTGGAATGGCGATTCTTGGCGTGTCTGTCTTCTGCTCTCTGGAGTTTGGCATCGCA 770
QY 872 TGTGCCACAGCCCATTTTCCCAACCA 897
DB 771 TCGGCATCTTCCCACTTTGGCTGCCA 796

RESULT 15

AAC99077
ID AAC99077 standard; cDNA; 1374 BP.

XX AAC99077;

XX 09-MAR-2001 (first entry)

XX	Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:305.
DE	Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
XX	detection; diagnosis; identification; cytostatic; neuroprotective;
KW	nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW	antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW	linkage analysis; tissue identification; tissue typing; forensic;
KW	neural; immune system; muscular; reproductive; gastrointestinal;
KW	muscular; cardiovascular; renal; proliferative; ss.
OS	Homo sapiens.
XX	WO200055320-A1.
PN	21-SEP-2000.
XX	08-MAR-2000; 2000WO-US05989.
PD	12-MAR-1999; 99US-0124270.
PF	(HUMA-) HUMAN GENOME SCI INC.
PP	Rosen CA, Ruben SM;
PR	WPI; 2000-579444/54.
XX	P-PSDB; AAB54312.
XX	New nucleic acid that is a pancreatic cancer antigen for preventing,
PT	treating, or ameliorating a medical condition, particular pancreatic
PT	cancer, or for use in assays for diagnosing a pathological condition -
XX	Claim 1; Page 731; 1379pp; English.
XX	AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC	proteins, called pancreatic cancer antigens, given in AAB54008 to
CC	AAB54466. The human pancreatic cancer antigens have cytostatic,
CC	neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC	gynaecological, cardiant and antiinflammatory activities, and can be used
CC	in gene therapy. The polynucleotide and proteins can be used for
CC	preventing, treating, or ameliorating a medical condition or in assays
CC	for diagnosing a pathological condition or a susceptibility to one in a
CC	subject. Binding partners to the proteins and the activity of the
CC	proteins can be identified. The pancreatic cancer antigens can be used to
CC	detect, treat or prevent pancreatic disorders, especially cancer.
CC	Agonists and antagonists to the antigens can be screened for. The
CC	pancreatic cancer antigen polynucleotides can be used to design nucleic
CC	acid hybridisation probes that can be used in chromosome mapping, linkage
CC	analysis, tissue identification and/or typing and a variety of forensic
CC	and diagnostic methods. The proteins can be used to generate antibodies
CC	which are used to purify, detect and target the polypeptides, including
CC	both in vivo and in vitro diagnostic and therapeutic methods. The
CC	proteins can be used to treat or prevent neural, immune system, muscular,
CC	reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC	proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC	sequences used in the exemplification of the present invention.
XX	Sequence 1374 BP; 361 A; 358 C; 301 G; 353 T; 1 other;
SQ	
Query Match	6.2%; Score 81.2; DB 21; Length 1374;
Best Local Similarity	58.3%; Pred. No. 2.4e-11;
Matches 190; Conservative	0; Mismatches 118; Indels 18; Gaps 2
QY	590 TCTACTGCTGTTATGTGGATACCAATCTGGGGTGGCCCTTCCTTTATTATCTCTGCG 649
Db	501 TCTATTTCATTCTACGGAGGCTTTCCCTCTTGGGGAGGCTTGTTGGTTTATCATTTTCAGA 560
QY	650 TCTCTCTCTGTGTACAGATCCAAAGAGCT---TTCGCCGTTGCTGTGGTGAAGAGCACCGTG 706
Db	561 TCTCTCTCTCGGAGCAGAGAAATACAGCCATATTCATTATTCCTGCTGCTGCGAGTTTG 620
QY	707 GGAATACAAATGTTAGTTCATTCTTGGCCCTTCATTGGAGTGATTCGTGCTGTGGTGGAT 766

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 08:51:42 ; Search time 1437.76 Seconds

(without alignments)

14835.236 Million cell updates/sec

Title: US-09-981-353-81

Perfect score: 1317

Sequence: 1 aeggaacaagtaagtccat.....ttttcttaataaaatgac 1317

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estl:*

9: gb_estl:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	728.2	55.3	859	13	BI760204 603045133
2	369.2	28.0	417	10	AW842562 MR2-CN003
3	368.2	28.0	415	10	AW842606 MR2-CN003
4	200.6	15.2	452	14	F15017 SSO4B08 Por
5	132.2	10.0	418	17	AQ213250 HS_3249_B
6	96.4	7.3	176	13	BG983751 MR4-CN014

ALIGNMENTS

RESULT 1
BI760204
LOCUS
DEFINITION 603045133f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185424 5',
859 bp mRNA linear EST 25-SEP-2001
mRNA sequence.
ACCESSION BI760204 GI:15751782
VERSION BI760204.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Primates; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 859)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11463 row: a column: 09
High quality sequence stop: 792.
Location/Qualifiers
1. .859

7 87.4 6.6 185 13 BG981760
8 79.6 6.0 880 13 BI766299
9 77.6 5.9 687 13 BM488591
10 75.6 5.7 673 10 BB629370
11 74.8 5.7 515 12 BG434337
12 74.8 5.7 594 10 AV717594
13 74.8 5.7 619 10 AV661743
14 74.8 5.7 680 12 BG546745
15 74.8 5.7 735 10 AV715678
16 74.8 5.7 869 13 BI766801
17 74.8 5.7 876 13 BI771044
18 74.8 5.7 934 13 BI821317
19 74.4 5.6 597 10 AV716310
20 74.4 5.6 860 9 AL551751
21 73.2 5.6 656 14 BM781186
22 71.2 5.4 481 9 AA160278
23 70.6 5.4 584 13 BI824549
24 70.6 5.4 1017 11 AK008099
25 69.6 5.3 853 12 BF574865
26 69.6 5.3 885 13 BI911270
27 69.2 5.3 619 17 BH338441
28 69 5.2 436 9 AI635515
29 69 5.2 540 9 AI552228
30 68.4 5.2 577 10 AV717053
31 67.8 5.1 773 9 AI263085
32 67.2 5.1 608 12 BF042446
33 67 5.1 528 12 BF868999
34 67 5.1 648 9 AI741604
35 66.4 5.0 449 9 AA469742
36 66.4 5.0 582 10 AV714513
37 65 4.9 504 12 BF522968
38 64.8 4.9 517 12 BG664654
39 64.6 4.9 886 13 BI908327
40 63.6 4.8 717 13 BI771406
41 63.4 4.8 483 13 BM430393
42 63.4 4.8 604 12 BF557830
43 63.2 4.8 491 12 BF556579
44 63.2 4.8 507 9 AI792758
45 63.2 4.8 686 10 AV733821

BG981760 MR3-CN014
BI766299 603052775
BM488591 pgm2n.pk0
BB629370 BB629370
BG434337 602506331
AV717594 AV717594
AV661743 AV661743
BG546745 602574115
AV715678 AV715678
BI766801 603056712
BI771044 603053393
BI821317 603038160
AV716310 AV716310
AL551751 AL551751
BM781186 MLN1_5_B1
AA160278 z087b11.r
BI824549 603035154
AK008099 Mus muscu
BF574865 602132285
BI911270 603062843
BH338441 CH230-74J
AI635515 ts95b02.x
AI552228 vf70b04.y
AV717053 AV717053
AI263085 qz35a05.x
BF042446 BP250003A
BF868999 OV3-ET010
AI741604 wq28d01.x
AA469742 vf70b04.r
AV714513 AV714513
BF522968 UI-R-C2p-
BG664654 DRABG07
BI908327 603066152
BI771406 603059433
BM430393 IDUO33F9.
BF557830 UI-R-C0-g
BF556579 UI-R-E1-f
AI792758 o101d12.y
AV733821 AV733821


```

Db      356 ATCCAAGAGCCTCCCGTGTCTGGTGAAGGAGCCCTGGGAATGAACATTGTTAGTTC 297
QY      727 TATCTTGGGCTTCATTTGGAGTGAATGCTGCTGGTGGGATATGTCATCAATGGGTAGC 786
Db      296 TATCTTGGGCTTCATTTGGAGTGAATGCTGCTGGTGGGATATGTCATCAATGGGTAGC 237
QY      787 TGGCCAAGACTACTGGGCGCTGCTTCTTGGAAGGCAATTTTCAGCCACGCTGATGATCTT 846
Db      236 TGGCCAAGACTACTGGGCGCTGCTTCTTGGAAGGCAATTTTCAGCCACGCTGATGATCTT 177
QY      847 CTCCTCTTGGAGTTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 905
Db      176 CTCCTCTTGGAGTTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 117
QY      906 CCACAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 965
Db      116 CCACAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57
QY      966 CAGCGTCTTCTTCAGCTCCTCCAGATGCAACAA 999
Db      56 CAGCGTCTTCTTCAGCTCCTCCAGATGCAACAA 23

RESULT 3
AW842606/c
LOCUS      MR2-CN0035-200300-203-h08 CN0035 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW842606
ACCESSION  AW842606.1 GI:7936589
VERSION     EST.
KEYWORDS    human.
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 415)
            Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsumura, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
            , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR2-CN0035-200
            300-203-h08&t3=2000-03-20&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 53
            High quality sequence stop: 415.
            Location/Qualifiers
            1. .415
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="CN0035"
            /dev_stage="Adult"
            /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent applich)
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the puc 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."

FEATURES
source
Location/Qualifiers
1. .415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0035"
/dev_stage="Adult"
/note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent applich)
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

BASE COUNT 124 a 93 c 112 g 86 t
ORIGIN
Query Match 28.0%; Score 368.2; DB 10; Length 415;
Best Local Similarity 97.7%; Pred. No. 1.3e-66;
Matches 384; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 609 GATACCCATCTCGGGTGGCCCTTCTTTTATTATCTCTGGCTCTCTCTCTCTCTCTCTCTCT 668
Db 415 GATACCCATCT- GGGTGGCCCTTCTTTTATTATCTCTGGCTCTCTCTCTCTCTCTCTCTCT 357
QY 669 CCAAGGAGCTTCCCGCTTCTGCTGGTGAAGGAGCGCTGGGAATGAACATTTGTTAGTCTTA 728
Db 356 CCAAGGAGCTTCCCGCTTCTGCTGGTGAAGGAGCGCTGGGAATGAACATTTGTTAGTCTTA 297
QY 729 TCTTGGCCCTTCATTTGGAGTGATTCGCTGCTGGTGGATATGTCATCAATGGGTAGCTG 788
Db 296 TCTTGGCCCTTCATTTGGAGTGATTCGCTGCTGGTGGATATGTCATCAATGGGTAGCTG 237
QY 789 GCCAAGACTACTGGGCCGTGCTTTCTGGAAAAGGCAATTTTCAGCCAGCGCTGATGATCTTCT 848
Db 236 TCCAGACTACTGGGCCGTGCGTCTCGGAAAAGGCAATTTTCAGCCAGCGCTGATGATCTTCT 177
QY 849 CCCTCTTGAGTTCCTTCGTAGCTTGTGCACAGCCCAATTTTGCACCAACCAACCAACCA 908
Db 176 CCCTCTTGAGTTCCTTCGTAGCTTGTGCACAGCCCAATTTTGCACCAACCAACCAACCA 117
QY 909 CAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968
Db 116 CAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57
QY 969 CGCTCTTCTTCAGCTCCTCCAGATGCAACAACT 1001
Db 56 CGCTCTTCTTCAGCTCCTCCAGATGCAACAACT 24

RESULT 4
F15017
LOCUS      SSO4B08 Porcine small intestine cDNA library Sus scrofa cDNA clone
DEFINITION O4B08, mRNA sequence.
ACCESSION  F15017
VERSION     F15017.1 GI:972439
KEYWORDS    EST.
SOURCE      pig.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
            1 (bases 1 to 452)
            Wintero, A.K., Fredholm, M. and Davies, W.
            Evaluation and characterization of a porcine small intestine cDNA
            library: analysis of 839 clones
            Mamm. Genome 7 (7), 509-517 (1996)
            96327607
            Contact: A.K. Wintero
            Department of Animal Science and Animal Health, Division of Animal
            Genetics, The Royal Veterinary and Agricultural University
            Bulowsvej 13, 1870 Frederiksberg C, Denmark.
            Location/Qualifiers
            1. .452
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone="O4B08"
            /clone_lib="Porcine small intestine cDNA library"
            /note="directionally cloned cDNA in XLI-blue MRF."

BASE COUNT 149 a 103 c 100 g 97 t 3 others
ORIGIN
Query Match 15.2%; Score 200.6; DB 14; Length 452;
Best Local Similarity 69.6%; Pred. No. 1e-31;
Matches 288; Conservative 0; Mismatches 117; Indels 9; Gaps 1;

QY 173 GAACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 232

```



```

RESULT 9
BM488591
LOCUS
DEFINITION
pgm2n.pk008.c22 Normalized Chicken Breast Muscle, EST 07-FEB-2002
Epiphyseal Growth Plate cDNA library (pgm2n) Gallus muscle, and
clone pgm2n.pk008.c22 5' similar to ref|NP_060186.1| (NM_017716)
membrane-spanning 4-domains, subfamily A, member 12 4-domains,
subfamily A, member 7 [Homo sapiens] dbj|BAA91020.1| (AK000224)
MS412 protein [Homo sapiens], mRNA sequence.
ACCESSION
BM488591
VERSION
BM488591.1 GI:18609522
SOURCE
EST.
ORGANISM
chicken.
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 687)
Codburn, L.A. and Monsonego-Ornan, E.
ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
Project
Unpublished (2002)
Contact: Larry A. Codburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: codburn@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
1. 687
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk008.c22"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast, leg: Embryo(d19); post-hatch(1d,1.3,5,7,9
,11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMDH108"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"
BASE COUNT 129 a 223 c 188 g 143 t 4 others
ORIGIN
Query Match 5.9%; Score 77.6; DB 13; Length 687;
Best Local Similarity 52.6%; Pred. No. 4.1e-06;
Matches 241; Conservative 0; Mismatches 184; Indels 33; Gaps 2;
QY 492 TAGGGGTATCCAGATCATGGTGGATTGATTCATCCATGGTTTGGAAATGTTTGTGT 551
Db 108 TGGGGACTATCCAGATCATGACCGGCTTCATGCATATGGCTTCGATGCGCTCCACCA 167
QY 552 TAATATCCCTCTCTTTAGAGAAGTATAGGTTTGGCTCTACTGCTGTTATGTTGGAT 611
Db 168 CACTCACC-----AATGCTACTCTCCATCTCATCCCGCGGAGA 209
QY 612 ACCCATTCGGGGTGGCCCTTTTATTTATCTCTGGCTCTCTCTGTTGTCAGCATCCA 671
Db 210 TCCCTTTCTGGCGGTGTGCTTTTCATCATCTCGGCTGCTGTCATTCGCGGAGAGA 269
QY 672 AGAGCTTTCCCTGCTGCTGGTGAAGGAGCGCTGGGATGAACATTTGTTAGTTCTATCT 731
Db 270 AGAGTCCACAGAGTGTGCGGTGAAGGGGAGCGGCCACGAACATCATGTCGTCATCT 329

```

```

QY 732 TGCGCTTCATTTGGAGTGAATTCCTGCTGGTGGATATGTCATCAATGGGTAGCTGG-- 789
Db 330 TTGCGCTCTCTGGGCATCGTCGCTTCATCATTTGACCTCAACTTCAACGGGCTGTACCGCT 389
QY 790 -----CCAAAGACTACTGCGCGTCTTTCTGGAAAGCATTTTCAGCACCGC 836
Db 390 CCAGCGATGACCTACTACAGTACCTTGTCTGCTCCAGGGAATGATCTCATCTCGTGC 449
QY 837 TCATGATCTCTTCCCTCTTGGAGTCTTCTGCTAGCTTGTGCCACAGCCCATTTTGCAACC 896
Db 450 TGCTCATCTTCCACCATCTGAGATTTCTGCTGCTGCCACTGCCAACTTCTGCTGTC 509
QY 897 AAGCAAAACACCAACCAATATGCTGCTCTGCTGTTATT 934
Db 510 GGGCAACAGCGCTCAGCCCAACGAGGCCATGCTGATT 547

RESULT 10
BM629370
LOCUS
DEFINITION
BB629370 RIKEN full-length enriched, adult male bone Mus musculus
CDNA clone 9830113C01 5', mRNA sequence.
ACCESSION
BB629370
VERSION
BB629370.1 GI:15399617
SOURCE
EST.
house mouse.
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 673)
Arawaka, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okada
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, F., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arawaka, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arawaka, T.,
Ishii, Y. and Hayashizaki, Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 594)
 Xu.X., Gu.J., Liu.F., Ou.J., Zhao.M., Li.Y., Huang.Q., Zhou.J., Song.H., Gu.Y., Yang.Y., Gao.G., Xiao.H., Li.N., Qian.B., Gao.X., Cheng.Z., Xu.S., Gu.W., Tu.Y., Jia.J., Fu.G., Ren.S., Zhong.M., Lu.G., Cheng.Z. and Han.Z.

TITLE

Homo sapiens cDNA DCB clones

JOURNAL

Unpublished (2000)

COMMENT

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..594

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DCBCE05"

/clone_lib="DCB"

/cell_type="dendritic cells"

/dev_stage="mature"

/lab_host="BM25.8"

/note="Vector: pTriplex2; Site_1: sflia; Site_2: sflib"

BASE COUNT 162 a 150 c 124 g 158 t

ORIGIN

Query Match 5.7%; Score 74.8; DB 10; Length 594;

Best Local Similarity 55.3%; Pred. No. 1.6e-05;

Matches 167; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 474 AAGAGAACCAAGCACTAGGGTGATCCACATCATGGTGGATGATGCACATGGTT 533

Db 244 ACGCAGAAATCAAGTATTGGGACTATCCAGATCTTGTGGCATGATGGTATTGAGCT 303

QY 534 TTGGAAATGTTTGTGTTAATATCTCTCTCTTTTAGAGAAGTATTAGTTTGGCTCTA 593

Db 304 TGGGGATCATTTGGCACTGCTCTCTCTC---CAAATTTTACCAGTGACTCTA 360

QY 594 CTGCTGTATTGGTGGATACCATCTCTGGGGTGGCTTTCTTTTATATCTCTGCTCTC 653

Db 361 CACTGTGAACCTCTGCTTACCATTCATAGGACCTTTTCTTTTATCATCTCTGCTCTC 420

QY 654 TCTCTGTGTGACATCCCAAGGAGCTTCCCGTGTCTGTGGAAGCAGCCCTGGGAATGA 713

Db 421 TATCAATGCCACACAGAAAGTTAACCAAGCTTTTGTGCATAGCAGCCCTGGTGGAA 480

QY 714 ACATTGTAGTTCTATCTTGGCTTCATGGAGTGATCTGCTGTGGTGGATATGTGCA 773

Db 481 GCATTCTGAGTCTGCTGCTGCCCTGGTGGTTCATTATCTCTGCTGCAACAGGCCA 540

QY 774 TC 775

Db 541 CC 542

RESULT 13

AV661743

LOCUS

AV661743 GLC Homo sapiens cDNA clone GLCXB07 3', mRNA sequence.

DEFINITION

AV661743

AV661743.1 GI:9882757

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 619)

Xu.X., Huang.J., Xu.Z., Qian.B., Zhu.Z., Yan.Q., Cai.T., Zhang.X.,

Xiao.H., Qu.J., Liu.F., Huang.Q., Cheng.Z., Li.N., Du.J., Hu.W.,

Shen.K., Lu.G., Fu.G., Zhong.M., Xu.S., Gu.W., Huang.W., Zhao.X., Hu.G., Gu.J., Chen.Z. and Han.Z.

Insight into hepatocellular carcinogenesis at transcriptome level

by comparing gene expression profiles of hepatocellular carcinoma

with those of corresponding noncancerous liver

Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

21625106

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..619

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="GLCXB07"

/clone_lib="GLC"

/tissue_type="corresponding non cancerous liver tissue"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 153 a 162 c 133 g 171 t

ORIGIN

Query Match 5.7%; Score 74.8; DB 10; Length 619;

Best Local Similarity 55.3%; Pred. No. 1.6e-05;

Matches 167; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 474 AAGAGAACCAAGCACTAGGGTGATCCACATCATGGTGGATGATGCACATGGTT 533

Db 177 ACGCAGAAATCAAGTATTGGGACTATCCAGATCTTGTGGCATGATGGTATTGAGCT 236

QY 534 TTGGAAATGTTTGTGTTAATATCTCTCTCTTTTAGAGAAGTATTAGTTTGGCTCTA 593

Db 237 TGGGGATCATTTGGCACTGCTCTCTCTCTCCAAATTTTACCCAAG---TGACTTCTA 293

QY 594 CTGCTGTATTGGTGGATACCATCTCTGGGGTGGCTTTCTTTTATCTCTGCTCTC 653

Db 294 CACTGTGAACCTCTGCTTACCATTCATAGGACCTTTTCTTTTATCATCTCTGCTCTC 353

QY 654 TCTCTGTGTGACATCCCAAGGAGCTTCCCGTGTCTGTGGAAGCAGCCCTGGGAATGA 713

Db 354 TATCAATGCCACACAGAAAGTTAACCAAGCTTTTGTGCATAGCAGCCCTGGTGGAA 413

QY 714 ACATTGTAGTTCTATCTTGGCTTCATGGAGTGATCTGCTGTGGTGGATATGTGCA 773

Db 414 GCATTCTGAGTCTGCTGCTGCCCTGGTGGTTCATTATCTCTGCTGCAACAGGCCA 473

QY 774 TC 775

Db 474 CC 475

RESULT 14

BG546745

LOCUS

602574115F1 NTH_MGC_77 Homo sapiens cDNA clone IMAGE:4702274 5',

mRNA sequence.

DEFINITION

BG546745

BG546745.1 GI:13545410

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 680)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI538 row: n column: 03
High quality sequence stop: 677.

FEATURES
source

1. 680
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4702274"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pONR-LIB (Clontech); Site_1:
SfiI (ggccgctcgccc); Site_2: SfiI (ggccattatgccc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 191 a 167 c 136 g 186 t
ORIGIN

Query Match 5.7%; Score 74.8; DB 12; Length 680;
Best Local Similarity 55.3%; Pred. No. 1.6e-05;
Matches 167; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 474 AAGAAGAACGAGCACTAGGGGTGATCCAGATCATGGTGGATTGATGCACATGGTT 533
DB 246 ACGGAGAAATCAAGTTATGGGACTATCCAGATCTTGTGGCATGATGTTAGCT 305
QY 534 TTGGAATTTGTTGTTTAAATATCCTCTCTCTTTAGAGAACTATTAGTTTGGCTCTA 593
DB 306 TGGGGATCATTTTGGCATCTGCTCTCTCTCCAAATTTTACCCAAG---TGACTTCTA 362
QY 594 CTGCTGTTATTTGGTGGATACCCATCTGCGGTGGCTTTCTTTTATATCTCTGCTCTC 653
DB 363 CACTGTGTAACCTCTGCTTACCCATCATAGGACCCCTTTTATATCATCTCTGCTCTC 422
QY 654 TCTCTGTGTCAGCATCCAGAGAGCTTCCCGTTGCTGTGTAAGGACGCTGGGAATGA 713
DB 423 TATCAATCCGACAGAGAAAGTTAAACCAAGCTTTTGTGTCATAGCAGCCCTGGTGGAA 482
QY 714 ACATGTTAGTTCTATCTTGGCTTCAATTTGGAGTGAATCTGCTGCTGGTGGATATGCA 773
DB 483 GCATTTGAGTGTCTGCTGCGCTGGTGGTTTCAATATCTGCTGCTCAACACAGGCCA 542
QY 774 TC 775
DB 543 CC 544

RESULT 15
AV715678
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

735 bp mRNA linear EST 11-OCT-2000
AV715678 DCB Homo sapiens cDNA clone DCBAXG06 5', mRNA sequence.
AV715678
AV715678.1 GI:10797195
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 735)

AUTHORS
TITLE
JOURNAL
COMMENT

Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203 P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source

1. 735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DCBAXG06"
/clone_lib="DCB"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/note="Vector: pRiplex2; Site_1: SfiIA; Site_2: SfiIB"
BASE COUNT 193 a 183 c 154 g 204 t
ORIGIN

Query Match 5.7%; Score 74.8; DB 10; Length 735;
Best Local Similarity 55.3%; Pred. No. 1.6e-05;
Matches 167; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

QY 474 AAGAAGAACGAGCACTAGGGGTGATCCAGATCATGGTGGATTGATGCACATGGTT 533
DB 236 ACGGAGAAATCAAGTTATTTGGGACTATCCAGATCTTGTGGCATGATGTTAGCT 295
QY 534 TTGGAATTTGTTGTTTAAATATCCTCTCTCTTTAGAGAACTATTAGTTTGGCTCTA 593
DB 296 TGGGGATCATTTTGGCATCTGCTCTCTCTCCAAATTTTACCCAAG---TGACTTCTA 352
QY 594 CTGCTGTTATTTGGTGGATACCCATCTGCGGTGGCTTTCTTTTATATCTCTGCTCTC 653
DB 353 CACTGTTGAACTCTGCTTACCCATCATAGGACCCCTTTTATATCATCTCTGCTCTC 412
QY 654 TCTCTGTGTCAGCATCCAGAGAGCTTCCCGTTGCTGTGTAAGGACGCTGGGAATGA 713
DB 413 TATCAATCCGACAGAGAAAGTTAAACCAAGCTTTTGTGTCATAGCAGCCCTGGTGGAA 472
QY 714 ACATGTTAGTTCTATCTTGGCTTCAATTTGGAGTGAATCTGCTGCTGGTGGATATGCA 773
DB 473 GCATTTGAGTGTCTGCTGCGCTGGTGGTTTCAATATCTGCTGCTCAACACAGGCCA 532
QY 774 TC 775
DB 533 CC 534

Search completed: July 16, 2003, 14:55:39
Job time : 1441.76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 10:35:27 ; Search time 51.3795 Seconds
(without alignments)
7860.974 Million cell updates/sec

Title: US-09-981-353-81
Perfect score: 1317
Sequence: 1 aaggaacaagtaagtccat.....ttttccttaataaatagtc 1317

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.8	5.7	945	4	US-09-149-476-168
2	71.2	5.4	2545	1	US-07-869-933-22
3	71.2	5.4	2545	4	US-09-103-663-22
4	53	4.0	1669	2	US-08-916-902A-2
5	53	4.0	1669	2	US-09-213-389-2
6	51.6	3.9	1661	1	US-08-318-492-3
7	51.6	3.9	1661	1	US-08-707-340-3
8	51.6	3.9	1661	2	US-08-994-578-3
9	51.6	3.9	7218	1	US-08-232-463-14
10	49.4	3.8	1308	4	US-09-724-864-15
11	44.4	3.4	7218	1	US-08-232-463-14
12	40.8	3.1	2621	2	US-08-553-619B-8
13	40.2	3.1	2065	4	US-09-370-473-5
14	38.6	2.9	1798	4	US-09-797-906-1
15	38.4	2.9	2207	6	5221620-3
16	38.4	2.9	2569	6	5221620-1
17	38.2	2.9	291	1	US-07-922-723A-7
18	38.2	2.9	291	1	US-07-799-828C-7
19	38.2	2.9	291	1	US-08-074-275-7
20	38.2	2.9	291	1	US-08-480-366-7
21	38.2	2.9	291	2	US-07-952-277A-7
22	38.2	2.9	152331	3	US-09-128-155-16
23	37.6	2.9	454	2	US-08-623-906A-6
24	37.2	2.8	227	2	US-08-520-678A-28
25	37.2	2.8	227	4	US-08-897-126-28
26	37.2	2.8	2674	4	US-09-817-180-1
27	36.8	2.8	2233	1	US-08-496-631-1

28	36.8	2.8	6638	2	US-08-070-301-2
29	36.6	2.8	6755	3	US-08-931-999-4
30	36.6	2.8	176373	3	US-09-128-155-17
31	36	2.7	1066	1	US-08-157-101A-4
32	36	2.7	1069	4	US-09-372-422A-7
33	36	2.7	3275	4	US-09-370-838-151
34	35.8	2.7	2447	2	US-09-014-969-14
35	35.6	2.7	356	2	US-08-520-678A-22
36	35.6	2.7	356	4	US-08-897-126-22
37	35.6	2.7	376	2	US-08-623-906A-18
38	35.6	2.7	9646	3	US-08-811-566-1
39	35.6	2.7	9646	4	US-09-034-756-1
40	35.4	2.7	1582	3	US-08-545-196B-10
41	35.4	2.7	1582	3	US-08-545-196B-12
42	35.4	2.7	2873	4	US-08-630-915A-193
43	35.4	2.7	12980	3	US-08-811-566-5
44	35.4	2.7	12980	4	US-09-034-756-5
45	35.2	2.7	1117	4	US-09-247-373B-33

ALIGNMENTS

RESULT 1
US-09-149-476-168
; Sequence 168, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23

1	EARLIER APPLICATION NUMBER: 60/047,500
2	EARLIER FILING DATE: 1997-05-23
3	EARLIER APPLICATION NUMBER: 60/047,587
4	EARLIER FILING DATE: 1997-05-23
5	EARLIER APPLICATION NUMBER: 60/047,492
6	EARLIER FILING DATE: 1997-05-23
7	EARLIER APPLICATION NUMBER: 60/047,598
8	EARLIER FILING DATE: 1997-05-23
9	EARLIER APPLICATION NUMBER: 60/047,613
10	EARLIER FILING DATE: 1997-05-23
11	EARLIER APPLICATION NUMBER: 60/047,582
12	EARLIER FILING DATE: 1997-05-23
13	EARLIER APPLICATION NUMBER: 60/047,596
14	EARLIER FILING DATE: 1997-05-23
15	EARLIER APPLICATION NUMBER: 60/047,612
16	EARLIER FILING DATE: 1997-05-23
17	EARLIER APPLICATION NUMBER: 60/047,632
18	EARLIER FILING DATE: 1997-05-23
19	EARLIER APPLICATION NUMBER: 60/047,607
20	EARLIER FILING DATE: 1997-05-23
21	EARLIER APPLICATION NUMBER: 60/043,314
22	EARLIER FILING DATE: 1997-04-11
23	EARLIER APPLICATION NUMBER: 60/043,569
24	EARLIER FILING DATE: 1997-04-11
25	EARLIER APPLICATION NUMBER: 60/043,311
26	EARLIER FILING DATE: 1997-04-11
27	EARLIER APPLICATION NUMBER: 60/043,671
28	EARLIER FILING DATE: 1997-04-11
29	EARLIER APPLICATION NUMBER: 60/043,674
30	EARLIER FILING DATE: 1997-04-11
31	EARLIER APPLICATION NUMBER: 60/043,669
32	EARLIER FILING DATE: 1997-04-11
33	EARLIER APPLICATION NUMBER: 60/043,312
34	EARLIER FILING DATE: 1997-04-11
35	EARLIER APPLICATION NUMBER: 60/043,313
36	EARLIER FILING DATE: 1997-04-11
37	EARLIER APPLICATION NUMBER: 60/043,672
38	EARLIER FILING DATE: 1997-04-11
39	EARLIER APPLICATION NUMBER: 60/043,315
40	EARLIER FILING DATE: 1997-04-11
41	EARLIER APPLICATION NUMBER: 60/048,974
42	EARLIER FILING DATE: 1997-06-06
43	EARLIER APPLICATION NUMBER: 60/056,886
44	EARLIER FILING DATE: 1997-08-22
45	EARLIER APPLICATION NUMBER: 60/056,877
46	EARLIER FILING DATE: 1997-08-22
47	EARLIER APPLICATION NUMBER: 60/056,889
48	EARLIER FILING DATE: 1997-08-22
49	EARLIER APPLICATION NUMBER: 60/056,893
50	EARLIER FILING DATE: 1997-08-22
51	EARLIER APPLICATION NUMBER: 60/056,630
52	EARLIER FILING DATE: 1997-08-22
53	EARLIER APPLICATION NUMBER: 60/056,878
54	EARLIER FILING DATE: 1997-08-22
55	EARLIER APPLICATION NUMBER: 60/056,662
56	EARLIER FILING DATE: 1997-08-22
57	EARLIER APPLICATION NUMBER: 60/056,872
58	EARLIER FILING DATE: 1997-08-22
59	EARLIER APPLICATION NUMBER: 60/056,882
60	EARLIER FILING DATE: 1997-08-22
61	EARLIER APPLICATION NUMBER: 60/056,637
62	EARLIER FILING DATE: 1997-08-22
63	EARLIER APPLICATION NUMBER: 60/056,903
64	EARLIER FILING DATE: 1997-08-22
65	EARLIER APPLICATION NUMBER: 60/056,888
66	EARLIER FILING DATE: 1997-08-22
67	EARLIER APPLICATION NUMBER: 60/056,879
68	EARLIER FILING DATE: 1997-08-22
69	EARLIER APPLICATION NUMBER: 60/056,880

1	EARLIER FILING DATE: 1997-08-22
2	EARLIER APPLICATION NUMBER: 60/056,894
3	EARLIER FILING DATE: 1997-08-22
4	EARLIER APPLICATION NUMBER: 60/056,911
5	EARLIER FILING DATE: 1997-08-22
6	EARLIER APPLICATION NUMBER: 60/056,636
7	EARLIER FILING DATE: 1997-08-22
8	EARLIER APPLICATION NUMBER: 60/056,874
9	EARLIER FILING DATE: 1997-08-22
10	EARLIER APPLICATION NUMBER: 60/056,910
11	EARLIER FILING DATE: 1997-08-22
12	EARLIER APPLICATION NUMBER: 60/056,864
13	EARLIER FILING DATE: 1997-08-22
14	EARLIER APPLICATION NUMBER: 60/056,631
15	EARLIER FILING DATE: 1997-08-22
16	EARLIER APPLICATION NUMBER: 60/056,845
17	EARLIER FILING DATE: 1997-08-22
18	EARLIER APPLICATION NUMBER: 60/056,892
19	EARLIER FILING DATE: 1997-08-22
20	EARLIER APPLICATION NUMBER: 60/057,761
21	EARLIER FILING DATE: 1997-08-22
22	EARLIER APPLICATION NUMBER: 60/047,595
23	EARLIER FILING DATE: 1997-05-23
24	EARLIER APPLICATION NUMBER: 60/047,599
25	EARLIER FILING DATE: 1997-05-23
26	EARLIER APPLICATION NUMBER: 60/047,588
27	EARLIER FILING DATE: 1997-05-23
28	EARLIER APPLICATION NUMBER: 60/047,585
29	EARLIER FILING DATE: 1997-05-23
30	EARLIER APPLICATION NUMBER: 60/047,586
31	EARLIER FILING DATE: 1997-05-23
32	EARLIER APPLICATION NUMBER: 60/047,590
33	EARLIER FILING DATE: 1997-05-23
34	EARLIER APPLICATION NUMBER: 60/047,594
35	EARLIER FILING DATE: 1997-05-23
36	EARLIER APPLICATION NUMBER: 60/047,589
37	EARLIER FILING DATE: 1997-05-23
38	EARLIER APPLICATION NUMBER: 60/047,593
39	EARLIER FILING DATE: 1997-05-23
40	EARLIER APPLICATION NUMBER: 60/047,614
41	EARLIER FILING DATE: 1997-05-23
42	EARLIER APPLICATION NUMBER: 60/043,578
43	EARLIER FILING DATE: 1997-04-11
44	EARLIER APPLICATION NUMBER: 60/043,576
45	EARLIER FILING DATE: 1997-04-11
46	EARLIER APPLICATION NUMBER: 60/047,501
47	EARLIER FILING DATE: 1997-05-23
48	EARLIER APPLICATION NUMBER: 60/043,670
49	EARLIER FILING DATE: 1997-04-11
50	EARLIER APPLICATION NUMBER: 60/056,632
51	EARLIER FILING DATE: 1997-08-22
52	EARLIER APPLICATION NUMBER: 60/056,664
53	EARLIER FILING DATE: 1997-08-22
54	EARLIER APPLICATION NUMBER: 60/056,876
55	EARLIER FILING DATE: 1997-08-22
56	EARLIER APPLICATION NUMBER: 60/056,881
57	EARLIER FILING DATE: 1997-08-22
58	EARLIER APPLICATION NUMBER: 60/056,909
59	EARLIER FILING DATE: 1997-08-22
60	EARLIER APPLICATION NUMBER: 60/056,875
61	EARLIER FILING DATE: 1997-08-22
62	EARLIER APPLICATION NUMBER: 60/056,862
63	EARLIER FILING DATE: 1997-08-22
64	EARLIER APPLICATION NUMBER: 60/048,964
65	EARLIER FILING DATE: 1997-06-06
66	EARLIER APPLICATION NUMBER: 60/056,887
67	EARLIER FILING DATE: 1997-08-22
68	EARLIER APPLICATION NUMBER: 60/056,908
69	EARLIER FILING DATE: 1997-08-22
70	EARLIER APPLICATION NUMBER: 60/048,964
71	EARLIER FILING DATE: 1997-06-06
72	EARLIER APPLICATION NUMBER: 60/057,650
73	EARLIER FILING DATE: 1997-09-05
74	EARLIER APPLICATION NUMBER: 60/056,884
75	EARLIER FILING DATE: 1997-08-22

Best Local Similarity 54.2%; Pred. NO. 1.4e-10;
Matches 167; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

QY 470 TTTAAGAGAGCAAGCAAGGCTAGGGGTGATCCAGATCATCTGGATTGATGACACATT 529
DB 214 TTGAAGAAGAGTTGGAGATCTCTGGCGTAACCCAAAGTTCTGGTTGATATGCGCTT 273
QY 530 GGTGTTGGAATGTTTGTGTTTAAATATCTCTCTTTTAGAGAAAGTATTAGGTTTGGC 589
DB 274 TGTGTTGGAACAGTTGCTGCTCCACACTCCAGACTTC--AGACTTTGAGCAGCAAGTG 330
QY 590 TCTACTGCTGTTATGTTGGATACCCATTCCTGGGTGCGCTTTCTTTTATTATCTCTGCG 649
DB 331 CTTTATTATATAGACAGGCTACCCATCTCTGGGTGCGAGTCTGTTGTTTGTCTGGA 390
QY 650 TCTCTCTCTGTCAGCATCAAGGAGCTTTCCCGTTCTCTGGTCAAGGCGAGCTGGGA 709
DB 391 TTTTGTCAATATGTCGCAAGGAAACACACTGTATCTGGTGAGAGGCGAGCTGGGA 450
QY 710 ATGAACATGTTAGTTCTATCTTGGCCCTTCATTGGAGTGATTCCTGCTGCTGGTGATG 769
DB 451 GCAACATGTCAGCAGCATCGCTGCAGGCTTGGGGATCGCCATATTGATCTCAATCTG 510
QY 770 TGCATCAA 777
DB 511 AGCAACAA 518

RESULT 4

US-08-916-902A-2

; Sequence 2, Application US/08916902A

; Patent No. 5871930

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E

; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/916,902A

; FILING DATE: Herewith

; CLASSIFICATION: 536

; PRIOR APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0371 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1669 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

LIBRARY: BRAINOT04
CLONE: 927955
US-08-916-902A-2

Query Match 4.0%; Score 53; DB 2; Length 1669;
Best Local Similarity 51.5%; Pred. NO. 1.9e-05;
Matches 153; Conservative 0; Mismatches 135; Indels 9; Gaps 1;

QY 439 AAATCCCAAGTCTGGCAACAGCAGTAATCAACTTTAAACAAGCAAGGCAAGGCTAGGGGT 498
DB 302 ACATCTGTGGAAGAGATTGCAAGAGAAGTCTTTGAAGGGAAGCAACCAAGTCTTGGGGT 361
QY 499 GATCCAGATCATGTTGGATTGATGACATTTGGTTTGGAAATTTGTTGTTGTTTAAATATC 558
DB 362 TGTGCAAGATCTGACTGCCCTGATGAGCCTTAGCATGGAATAACAATGATGTGTATGCG 421
QY 559 CTTCTCTTTTAGAAGAATATTAGGTTTGGCTCTACTGCTGTATTGTTGGTATACCCATT 618
DB 422 ATCTAATACT-----TATGGAAGTAACCCCTATTTCCTGTATATCGGGTACACAAT 472
QY 619 CTGGGTGGCCTTTCTTTTATTATCTCTGGCTCTCTCTCTGTCAGCATCCAAAGGAGCT 678
DB 473 TTGGGGTCAAGTAATGTTTATTATTTTTCAGGATCTTGTCAATTCGACGAGGAATTAGAAC 532
QY 679 TTCCCGTTGCTGTTGGAAGGAGCGCTGGGAATGAACAATTTAGTTCTATCTTGGC 735
DB 533 TACAAAAGGCGCTGGTCCGAGGTAGTCTAGGAATGAATATACACAGCTCTGTACTGCG 589

RESULT 5

US-09-213-389-2

; Sequence 2, Application US/09213389

; Patent No. 5977072

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E

; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/213,389

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/916,902

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0371 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1669 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-09-213-389-2

Query Match 4.08; Score 53; DB 2; Length 1669;
Best Local Similarity 51.58; Pred. No. 1.9e-05; Indels 9; Gaps 1;
Matches 153; Conservative 0; Mismatches 135; Indels 9; Gaps 1;

QY 439 AAATCCAAAGTGTGGGAACAGCAATGAACATTTAAAGAAAGCAAGCACTAGGGGT 498
DB 302 ACATCTGTGAAAGGATTGCAAGAGAGAGTTCTTGAAGGGAGAACCCAAAGTCTTGGGT 361
QY 499 GATCCAGATCATGTTGGATTTGATGACATGTTGTTTGGATGTTTGTGTTTAATATC 558
DB 362 TGTGACAGATTCTGACTCCCTCGATGAGCCCTAGATGGGAATAACAATGATGTATGGC 421
QY 559 CTTCCTCTTTAGAGAAGTATTAGGTTTGGCTCTACTGCTGTTTATTGTTGGTATACCCATT 618
DB 422 ATCTAATACT-----TATGGAAGTAACCCCTATTTCGGTGATATCGGGTACACAAT 472
QY 619 CTGGGTGGGCTTTCTTTTATTATCTCTGGCTCTCTCTCTGTCAGATCCCAAGGAGCT 678
DB 473 TTGGGGTGCAGTAATGTTTATTATTTCAGGATCCTTGTCAATTCGAGCAGGAATTAGAAC 532
QY 679 TTCCCGTGTCTGTGTAAGGAGCGCTGGGATGAACATTTGTTCTATCTTTGGC 735
DB 533 TACAAAAGGCGCTGTCGAGGAGTAGCTAGGAATGAATATACCAAGCTCTGTACTGGC 589

RESULT 6
US-08-318-492-3
; Sequence 3, Application US/08318492
; Patent No. 5552312
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,492
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BIH94-03
; TELEPHONE: (617) 861-9540
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

LOCATION: 97..741
US-08-318-492-3

Query Match 3.98; Score 51.6; DB 1; Length 1661;
Best Local Similarity 48.08; Pred. No. 4.9e-05; Indels 0; Gaps 0;
Matches 147; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 473 AAAGAAGCAAGCAAGGCACTAGGGGTGATCCAGATCATGTTGGATTGATGCACATGGT 532
DB 229 AAGCAAAATACAAAGTTCTTGGGGCCATCCAGATCCTGAATGCAGCAATGATTCGGCT 288
QY 533 TTTGGAATGTTTGTGTTTAATATCCTTCTCTTTTATAGAAAGTATTAGGTTTTCCTCT 592
DB 289 TTGGGTCTCTTCTGGGTCTCTTGCATATACCCATACCACTTCCAAAAGCACTTCTTTTC 348
QY 593 ACTGCTCTTATGTTGATATACCCATCTGGGGTGGCCCTTTCTTTTATTATCTCTGGCTCT 652
DB 349 TTCACCTTCTACACAGGCTACCCGATTTGGGGTCTGTGTTTCTGTAGTTCAGGAACC 408
QY 653 CTCTCTCTGTCAGCATCCAGGAGCTTTCCCGTTGTTCTGGTGAAGGAGGCTGGGAATG 712
DB 409 TTGCTCTGTGTAGCAGGATAAACCCACAAAGAACATGGATACAGAACAGTTTGGAAATG 468
QY 713 AACATTGTTAGTTCTATCTTGGCCCTTCATTGGAGTGATTCCTGCTGCTGATATGTC 772
DB 469 AACATTGCCAGTGTACAAATTTGCATAGTGGGAGTCTCTTCTCTCAATAATATAGCA 528
QY 773 ATCAAT 778
DB 529 GTTAAT 534

RESULT 7
US-08-707-340-3
; Sequence 3, Application US/08707340
; Patent No. 5705615
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,340
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,492
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/675,648
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BIH94-03A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 3:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..738
;
US-08-707-340-3

Query Match          3.9%; Score 51.6; DB 1; Length 1661;
Best Local Similarity 48.0%; Pred. No. 4.9e-05;
Matches 147; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

Qy 473 AAAGAAGCAAGAGGCACTAGGCGTATCCAGATCATGTTGATTCACATTTGGT 532
Db 229 AAAGCAAAATACAAAGTTCTTGGGGCCATCCAGATCCTGAATGCAGCAATGATTCGGCT 288
Qy 533 TTGGGAATTTGTTGTTTAATATCTCTTTTAGAGAAATATTAGTTTTCCTCT 592
Db 289 TTGGGTCTCTTCTGGGTCTCTTGCATACCCATACCCATACCCATCTCTTTTC 348
Qy 593 ACTCCTGTTATGTTGATACCCATCTGGGGTGGCCCTTTCTTTTATTATCTCTGGCTCT 652
Db 349 TTCACCTTCTACACAGGCTACCCGATTTGGGGTCTGTTCTGTAGTTTCAGGAACC 408
Qy 653 CTCCTCTGTGTCAGCATCAAGGAGCTTTCCGTTGTTCTGTGTAAGGCAAGCCTGGGAATG 712
Db 409 TTGCTGTTGAGCAGGGATTAACCCACCAAGAACATGGATACAGAACAGTTTGGGAATG 468
Qy 713 AACATTGTTAGTTCTATCTTTGGCCCTTCATTGGAGTGATTCGTGCTGGTGGATATGTC 772
Db 469 AACATGCCAGTGTACAAATTCGACTAGTGGGGACTGCTTTCTCTCACTAAATATAGCA 528
Qy 773 ATCAAT 778
Db 529 GTTAAT 534

RESULT 8
US-08-994-578-3
; Sequence 3, Application US/08994578
; Patent No. 5972688
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lellias, Jean-Michel
; TITLE OF INVENTION: Htm4 METHODS OF TREATMENT AND ASSAYS,
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994.578
; FILING DATE: December 19, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/707.340
; FILING DATE: 03-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/675.648
; FILING DATE: 03-JUL-1996

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318.492
; FILING DATE: 06-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BIH94-03A22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..738
;
US-08-994-578-3

Query Match          3.9%; Score 51.6; DB 2; Length 1661;
Best Local Similarity 48.0%; Pred. No. 4.9e-05;
Matches 147; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

Qy 473 AAAGAAGCAAGAGGCACTAGGCGTATCCAGATCATGTTGATTCACATTTGGT 532
Db 229 AAAGCAAAATACAAAGTTCTTGGGGCCATCCAGATCCTGAATGCAGCAATGATTCGGCT 288
Qy 533 TTGGGAATTTGTTGTTTAATATCTCTTTTAGAGAAATATTAGTTTTCCTCT 592
Db 289 TTGGGTCTCTTCTGGGTCTCTTGCATACCCATACCCATACCCATCTCTTTTC 348
Qy 593 ACTCCTGTTATGTTGATACCCATCTGGGGTGGCCCTTTCTTTTATTATCTCTGGCTCT 652
Db 349 TTCACCTTCTACACAGGCTACCCGATTTGGGGTCTGTTCTGTAGTTTCAGGAACC 408
Qy 653 CTCCTCTGTGTCAGCATCAAGGAGCTTTCCGTTGTTCTGTGTAAGGCAAGCCTGGGAATG 712
Db 409 TTGCTGTTGAGCAGGGATTAACCCACCAAGAACATGGATACAGAACAGTTTGGGAATG 468
Qy 713 AACATTGTTAGTTCTATCTTTGGCCCTTCATTGGAGTGATTCGTGCTGGTGGATATGTC 772
Db 469 AACATGCCAGTGTACAAATTCGACTAGTGGGGACTGCTTTCTCTCACTAAATATAGCA 528
Qy 773 ATCAAT 778
Db 529 GTTAAT 534

RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/232,463
FILING DATE: 26-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: EP 91 114 300.6
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 3.9%; Score 51.6; DB 1; Length 7218;
Best Local Similarity 3.1%; Pred. No. 0.00011;
Matches 12; Conservative 221; Mismatches 155; Indels 0; Gaps 0;

QY 115 AGACAAGTCTTGACCTATGTTGAGCGCAGTTGAAATAATGAGGAGAGATAAAACCATGA 174
DB 1451 ATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1392

QY 175 ACAGAAACAAGAAAGAAACAGAGAGAGAAATGAAAAGACATAATGATGTCAATCCAA 234
DB 1391 RRR 1332

QY 235 GCAACAAGCATGCTGAAGTAATGAACCATACCAACCCCTTACCACCAAGCAGCTT 294
DB 1331 RRR 1272

QY 295 TATGGCTCTGGGATTTCAACAGCCTCTGGGTTCAACTAGAAAACCAAGCTCAGGG 354
DB 1271 RRR 1212

QY 355 TGCTCAGCGTGTACGCGCTTACGCGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAAC 414
DB 1211 RRR 1152

QY 415 GGTCAAGGAATATACAAATGATAATCAAGTGTGGGACAGCAGTAAATGAACITTA 474
DB 1151 RRR 1092

QY 475 AGAAGAACAAAGCAGCTAGGGTGTATC 502
DB 1091 RRR 1064

RESULT 10
US-09-724-864-15
Sequence 15, Application US/09724864
Patent No. 6380362
GENERAL INFORMATION:
APPLICANT: Watson, James D
APPLICANT: Murison, James G.
TITLE OF INVENTION: Polynucleotides, polypeptides expressed
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
FILE REFERENCE: 11000.1050U
CURRENT APPLICATION NUMBER: US/09/724,864
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
PRIOR FILING DATE: 1999-12-23

NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 1308
TYPE: DNA
ORGANISM: Mouse
US-09-724-864-15

Query Match 3.8%; Score 49.4; DB 4; Length 1308;
Best Local Similarity 50.2%; Pred. No. 0.00018;
Matches 150; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 471 TTAAGAAGAAAGCAAGCACTAGGGGTGATCCAGATCATGTTGATGATGATGATGATG 530
DB 186 TAAAGCTGAGATCAAGTGATGGCGCAATCCAGATCATGTTGCTGTGATGTTGCA 245

QY 531 GTTTTGGAAATGTTTGTCTTTAATATCTCTCTTTTATAGAGAAATGATGATGTTTGCCT 590
DB 246 GTCTGGGAATCATTTTGGCATCTGTTCCCTCCAACTACACTTTACCTCAGTGTTT--T 302

QY 591 CTACTGCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 650
DB 303 CCATCTGTTAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362

QY 651 CTCTCTCTGTCAGCATCAAGGAGCTTTCCCGTTGCTGTTGTTGTTGTTGTTGTTGTTG 710
DB 363 TTTCTGCTATGTCACAGAGAAAAGATGACTAAGCCTTTGGTTACAGCAGCCTAGCCC 422

QY 711 TGAACATGTTAGTTCTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769
DB 423 TGAGCATCTGAGTGTCT 481

RESULT 11
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFELINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 13:28:07 ; Search time 173.297 Seconds
(without alignments)
12133.023 Million cell updates/sec

Title: US-09-981-353-81
Perfect score: 1317
Sequence: 1 aaggacaagaagtaagtcacat.....tttttccttaataaagtc 1317

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1114047 seqs, 798260406 residues
Total number of hits satisfying chosen parameters: 2228094

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1317	100.0	1317	9	US-09-981-353-81
2	1317	100.0	1318	9	US-10-158-646-77
3	1084.4	82.3	1291	9	US-10-054-988-17
4	1084.4	82.3	1291	10	US-09-739-254-17
5	1084.4	82.3	1291	10	US-09-904-615-17
6	1079	81.9	1292	10	US-09-904-615-61
7	1079	81.9	1292	10	US-09-739-254-61
8	1079	81.9	1292	10	US-09-904-615-61
9	389.6	29.6	441	9	US-09-925-299-727
10	389.6	29.6	441	10	US-09-925-299-727
11	303.6	23.1	351	9	US-09-803-719-494
12	292.6	22.2	342	9	US-09-803-719-638
13	288.6	21.9	383	9	US-09-803-719-481
14	92.6	7.0	125	10	US-09-563-817-780
15	81.2	6.2	753	9	US-10-017-754-2003
16	81.2	6.2	1236	9	US-10-156-136-5
17	81.2	6.2	1236	9	US-10-156-136-51
18	81.2	6.2	1336	9	US-09-736-457-1676
19	81.2	6.2	1336	9	US-09-902-941-1676

20	81.2	6.2	1336	9	US-09-849-626-1676	Sequence 1676, Ap
21	81.2	6.2	1336	9	US-10-017-754-1676	Sequence 1676, Ap
22	81.2	6.2	1374	10	US-09-925-297-305	Sequence 305, App
23	79.6	6.0	861	9	US-09-902-941-1877	Sequence 1877, Ap
24	79.6	6.0	861	9	US-09-849-626-1877	Sequence 1877, Ap
25	79.6	6.0	861	9	US-10-017-754-1877	Sequence 1877, Ap
26	79.6	6.0	1155	9	US-09-902-941-1875	Sequence 1875, Ap
27	79.6	6.0	1155	9	US-09-849-626-1875	Sequence 1875, Ap
28	79.6	6.0	1155	9	US-10-017-754-1875	Sequence 1875, Ap
29	79.6	6.0	1353	9	US-09-902-941-1873	Sequence 1873, Ap
30	79.6	6.0	1353	9	US-09-849-626-1873	Sequence 1873, Ap
31	79.6	6.0	1353	9	US-10-017-754-1873	Sequence 1873, Ap
32	74.8	5.7	932	9	US-10-220-946-7	Sequence 7, Appl
33	74.8	5.7	945	9	US-09-809-391-168	Sequence 168, App
34	74.8	5.7	1036	9	US-10-220-946-5	Sequence 5, Appl
35	74.8	5.7	1728	10	US-09-822-849A-23	Sequence 23, Appl
36	74.8	5.7	1743	10	US-09-925-302-314	Sequence 314, App
37	74.8	5.7	1871	9	US-10-106-698-2116	Sequence 2116, Ap
38	60	4.6	486	9	US-10-046-935-2044	Sequence 2044, Ap
39	60	4.6	486	9	US-09-878-178-2044	Sequence 2044, Ap
40	60	4.6	486	9	US-10-146-502-2044	Sequence 2044, Ap
41	53.8	4.1	1146	9	US-10-207-655-130	Sequence 130, App
42	53	4.0	1330	9	US-09-374-046A-151	Sequence 151, App
43	53	4.0	2438	9	US-09-822-846-227	Sequence 227, Appl
44	52.4	4.0	747	10	US-09-731-872-17	Sequence 17, Appl
45	52.4	4.0	760	10	US-09-821-821-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-981-353-81
; Sequence 81, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 81
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 611082CB1
US-09-981-353-81

Query Match	100.0%	Score	1317	DB	9	Length	1317
Best Local Similarity	100.0%	Pred. No.	0				
Matches 1317	Conservative	0	Mismatches	0	Indels	0	Gaps
0							
Qy	1	AAGGAACAAGTAAGTCCATTGATAGCTTCTTTCCTATCTCTCTCCAAATCAATGGCA	60				
Db	1	AAGGAACAAGTAAGTCCATTGATAGCTTCTTTCCTATCTCTCTCCAAATCAATGGCA	60				
Qy	61	CRAACTGTGGCTGGTCTACCTGTGTGGGTCTCTTAGATTGAGGGATGAAGCA	120				
Db	61	CRAACTGTGGCTGGTCTACCTGTGTGGGTCTCTTAGATTGAGGGATGAAGCA	120				
Qy	121	GTTCCTTACTCTATGTGTAGGCCAGTTGAAATGAGGGAGATAAACCATCAAGCA	180				
Db	121	GTTCCTTACTCTATGTGTAGGCCAGTTGAAATGAGGGAGATAAACCATCAAGCA	180				
Qy	181	CAAGAAGAAACAAAACAGAGAGGAATGAAAAGACATATGATGTCTATCCAGCAAC	240				
Db	181	CAAGAAGAAACAAAACAGAGAGGAATGAAAAGACATATGATGTCTATCCAGCAAC	240				


```

|||||
662 GTGAGCATCAAGAGCTTTCCCGTTGCTGGTGAAGGAGCGCTGGGAATGAACATGT 721
QY 721 TAGTTCTATCTTGGCCTTCAATGGAGTGATTCCTGCTGCTGGTGAATATGTCATCAATGG 780
Db 722 TAGTTCTATCTTGGCCTTCAATGGAGTGATTCCTGCTGCTGGTGAATATGTCATCAATGG 781
QY 781 GGTAGCTGGCCAAAGACTACTGGGCGCTTTCTGGAAAGGAGATTCAGCCAGCTGAT 840
Db 782 GGTAGCTGGCCAAAGACTACTGGGCGCTTTCTGGAAAGGAGATTCAGCCAGCTGAT 841
QY 841 GATCTTCTCCCTCTTGGAGTTCTTCTGTPAGCTTGTGCCACAGCCCATTTTGGCCAAACCAAGC 900
Db 842 GATCTTCTCCCTCTTGGAGTTCTTCTGTPAGCTTGTGCCACAGCCCATTTTGGCCAAACCAAGC 901
QY 901 AAACACACACCAACCAATATGCTGCTGCTGGTTATTTCCAAATATGATGAAAGCAACCCCTGT 960
Db 902 AAACACACACCAACCAATATGCTGCTGCTGGTTATTTCCAAATATGATGAAAGCAACCCCTGT 961
QY 961 GACACGAGCTTCTTTCAGCTCTCCAGATGCAACCACTACTCAGCTAATGCCCTAA 1020
Db 962 GACACGAGCTTCTTTCAGCTCTCCAGATGCAACCACTACTCAGCTAATGCCCTAA 1021
QY 1021 ATAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACACTACTTGCAAAACCT 1080
Db 1022 ATAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACACTACTTGCAAAACCT 1081
QY 1081 TCTTAAAGAAGATGCTTTTATGCTACATGATTTCTAGTCTTTAAACACTGTGTGA 1140
Db 1082 TCTTAAAGAAGATGCTTTTATGCTACATGATTTCTAGTCTTTAAACACTGTGTGA 1141
QY 1141 GATTTGTTTTAGTGTGCTCAATGATGCTGATCTCCCTCAGCTCTCTTCTCTAC 1200
Db 1142 GATTTGTTTTAGTGTGCTCAATGATGCTGATCTCCCTCAGCTCTCTTCTCTAC 1201
QY 1201 ATTACCACTACTACATGCTGCAAGAGTGAAGGATCAGAGACTGAAAATGATTTCTGCA 1260
Db 1202 ATTACCACTACTACATGCTGCAAGAGTGAAGGATCAGAGACTGAAAATGATTTCTGCA 1261
QY 1261 ACTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTTCTTTTCCCTTAAATAAGTGC 1317
Db 1262 ACTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTTCTTTTCCCTTAAATAAGTGC 1318
```

```

RESULT 3
US-10-054-988-17
; Sequence 17, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032PI
; CURRENT APPLICATION NUMBER: US/10/054,988
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1279)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
```

```

; LOCATION: (1286)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a.t.g, or c
US-10-054-988-17

Query Match      82.3%  Score 1084.4;  DB 9;  Length 1291;
Best Local Similarity 97.8%  Pred. No. 6.2e-297;
Matches 1111;  Conservative 1;  Mismatches 17;  Indels 7;  Gaps 1;

QY 182 AAGAAAGAAACAAACAGAGAGATGAAAAGACATAATATGATGTCATCAAGCCCAACA 241
Db 75 AAAGAGAAACATAGAGGTGCCAAAGGAACAAAGACATAATATGATGTCATCAAGCCCAACA 134
QY 242 AGCCATGCTCAAGTAAATGAAACCATACCAACCCCTTACCACCAAGCAGCTTTATGGCT 301
Db 135 AGCCATGCTCAAGTAAATGAAACCATACCAACCCCTTACCACCAAGCAGCTTTATGGCT 194
QY 302 CCTGGATTTCAACAGCCCTCTGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTCTCAG 361
Db 195 CCTGGATTTCAACAGCCCTCTGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTCTCAG 254
QY 362 CGTCTCAGCCCTACGGCATCACATCTCCGGGAATCTTGTAGCAGTCAACCGGGTCAA 421
Db 255 CGTCTCAGCCCTACGGCATCACATCTCCGGGAATCTTGTAGCAGTCAACCGGGTCAA 314
QY 422 GGAATATACAAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTAAAGAGAA 481
Db 315 GGAATATACAAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTAAAGAGAA 374
QY 482 GCAAGGCACATAGGGGTGATCCAGATCATGTTGGATGATGACATTTGGTTTGGAAATT 541
Db 375 GCAAGGCACATAGGGGTGATCCAGATCATGTTGGATGATGACATTTGGTTTGGAAATT 434
QY 542 GTTTTGTGTTTAATATCTTCTTTTAGAGAAGTATTAGGTTTGGCTCTACTCTCTGT 601
Db 435 GTTTTGTGTTTAATATCTTCTTTTAGAGAAGTATTAGGTTTGGCTCTACTCTCTGT 494
QY 602 ATTTGTGATACCCATCTTGGGGTGGCTTTCTTTTATTATCTCTGGCTCTCTCTCTGT 661
Db 495 ATTTGTGATACCCATCTTGGGGTGGCTTTCTTTTATTATCTCTGGCTCTCTCTCTGT 554
QY 662 TCAGCATCAAAGGAGCTTTCGGTGTGGAAGGAGCCTGGGAATGAACATTTGT 721
Db 555 TCAGCATCAAAGGAGCTTTCGGTGTGGAAGGAGCCTGGGAATGAACATTTGT 614
QY 722 AGTTCTATCTTGGCCCTTCATTTGAGTGTGCTGCTGGTGGATATGTGCATCAATGG 781
Db 615 AGTTCTATCTTGGCCCTTCATTTGAGTGTGCTGCTGGTGGATATGTGCATCAATGG 674
QY 782 GTAGCTGGCCAAAGACTACTGGCCCGTCTTCTGGAAGAGGATTTTCCAGCCAGCTGATG 841
Db 675 GTAGCTGGCCAAAGACTACTGGCCCGTCTTCTGGAAGAGGATTTTCCAGCCAGCTGATG 734
QY 842 ATCTTCTCCCTCTTGGAGTCTTCTGAGCTTGTGCCACAGCCCATTTTCCCAACCAAGCA 901
Db 735 ATCTTCTCCCTCTTGGAGTCTTCTGAGCTTGTGCCACAGCCCATTTTCCCAACCAAGCA 794
QY 902 AACACCAACCAATATGCTGCTGCTGTTATTTCCAAATATGATGAAAGCAACCCCTGTG 961
Db 795 AACACCAACCAATATGCTGCTGCTGTTATTTCCAAATATGATGAAAGCAACCCCTGTG 854
QY 962 ACACGAGCTCTTCTTTCAGCTCTCTCCAGATGCAACCACTACTCAGCTAATGCCCTAA 1021
Db 855 ACACGAGCTCTTCTTTCAGCTCTCTCCAGATGCAACCACTACTCAGCTAATGCCCTAA 910
QY 1022 TAGTAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACACTACTTGCAAAACCT 1081
Db 911 ---TAAAGAAAAGGGGTATCAGTCTAATCTCATGGAGAAAACACTACTTGCAAAACCT 967
QY 1082 CTTAAGCAAGATGCTTTTATTGTTCTACAATGATTTCTTAGTCTTTAAAAACTGTGTTGAG 1141
Db 1141 CTTAAGCAAGATGCTTTTATTGTTCTACAATGATTTCTTAGTCTTTAAAAACTGTGTTGAG 1141
```

Db 968 CTTAAGAAGATGCTCTTTATTATGTTCTACAAATGATTTCTAGTCTTTAAACAACTGTGTTGAG 1027
Qy 1142 ATTTGTTTTAGTTGGTCGCTAAATGATGCTGTATCTCCCTTCACTGCTCTCTCTTCTTACA 1201
Db 1028 ATTTGTTTTAGTTGGTCGCTAAATGATGCTGTATCTCCCTTCACTGCTCTCTTCTTCTTACA 1087
Qy 1202 TTACCACTACTACATGCTGCGCAAGGTAAGGATCAGAGGACTGAAACAAATGATTTCTGCAA 1261
Db 1088 TTACCACTACTACATGCTGCGCAAGGTAAGGATCAGAGGACTGAAACAAATGATTTCTGCAA 1147
Qy 1262 CTCCTTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCCCTTAATAAATGTC 1317
Db 1148 CTCCTTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCCCTTAATAAATGTC 1203

RESULT 4

US-09-739-254-17
; Sequence 17, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT FILING DATE: 2000-12-19
; EARLIER FILING DATE: 2000-02-23
; EARLIER FILING DATE: 2000-02-23
; EARLIER FILING DATE: 1999-08-24
; EARLIER FILING DATE: 1999-08-24
; EARLIER FILING DATE: 1998-08-25
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1279)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1286)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-739-254-17

Query Match 82.3%; Score 1084.4; DB 10; Length 1291;
Best Local Similarity 97.8%; Pred. No. 6.2e-297;
Matches 1111; Conservative 1; Mismatches 17; Indels 7; Gaps 1;
Qy 182 AAGAAGAAACAAACAGAGGAGGAAATGAAAGACATAATGATGTCATCCAAAGCCAAACA 241
Db 75 AAAGAGGAACATAGAGTGCCAAAGGAACAAAGACATAATGATGTCATCCAAAGCCAAACA 134
Qy 242 AGCCATGCTGAAGTAAATGAACACCATACCCACCCCTTACCACCAAGCAGCTTTATGCT 301
Db 135 AGCCATGCTGAAGTAAATGAACACCATACCCACCCCTTACCACCAAGCAGCTTTATGCT 194
Qy 302 CCGATTTTCAACACCCCTTGGGTTCAATCACTTAGAAACCAAGCTCAGGGTCTCAG 361
Db 195 CCGATTTTCAACACCCCTTGGGTTCAATCACTTAGAAACCAAGCTCAGGGTCTCAG 254
Qy 362 CCGTCTCAGCCCTACCGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCCGGGTCAA 421
Db 255 CGTCTCAGCCCTACCGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCCGGGTCAA 314

Qy 422 GGAATATACAAATGATAAAATCCAAAGTGTGGGAACAGCAGTAATGAACATTTTAAAGAGAA 481
Db 315 GGAATATACAAATGATAAAATCCAAAGTGTGGGAACAGCAGTAATGAACATTTTAAAGAGAA 374
Qy 482 GCAAGGCACTAGGGTGATCCAGATCATGCTGGATTGATGACATTTGGTTTGGGAATT 541
Db 375 GCAAGGCACTAGGGTGATCCAGATCATGCTGGATTGATGACATTTGGTTTGGGAATT 434
Qy 542 GTTTTGTCTTTTAAATATCTCTCTTTTAAAGAAATATTAGTCTTTTGGCTCTACTGCTCTT 601
Db 435 GTTTTGTCTTTTAAATATCTCTCTTTTAAAGAAATATTAGTCTTTTGGCTCTACTGCTCTT 494
Qy 602 ATGCTGGATACCCATTTCTGGGTGGCCTTTCTTTTATTTATCTCTGGCTCTCTCTCTGTG 661
Db 495 ATGCTGGATACCCATTTCTGGGTGGCCTTTCTTTTATTTATCTCTGGCTCTCTCTCTGTG 554
Qy 662 TCAGATCCAAAGGAGCTTTCCCGTTGCTGCTGTAAGGCAAGCCCTGGGAATGAACATTTGTT 721
Db 555 TCAGATCCAAAGGAGCTTTCCCGTTGCTGCTGTAAGGCAAGCCCTGGGAATGAACATTTGTT 614
Qy 722 AGTTCTATCTTTGGCCTTCATTGGAGTGATTTCTGCTGCTGGTGATATGTCATCAATGGG 781
Db 615 AGTTCTATCTTTGGCCTTCATTGGAGTGATTTCTGCTGCTGGTGATATGTCATCAATGGG 674
Qy 782 GTAGCTGGCCAAAGACTACTGGCCGCTGCTTTCTGAAAGAGCATTTTTCAGCCACGCTGATG 841
Db 675 GTAGCTGGCCAAAGACTACTGGCCGCTGCTTTCTGAAAGAGCATTTTTCAGCCACGCTGATG 734
Qy 842 ATCTTCTCCCTTTGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
Db 735 ATCTTCTCCCTTTGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 794
Qy 902 AACCCACAAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Db 795 AACCCACAAACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
Qy 962 ACACAGGCTCTTCTTACGCTCTCCAGATGCAACACTACTCAGCTAATGCTGCTGCTGCTTAA 1021
Db 855 ACACAGGCTCTTCTTACGCTCTCCAGATGCAACACTACTCAGCTAATGCTGCTGCTGCTTAA 910
Qy 1022 TAGTAAAGAAAGAGGGGTATCAGTCTAATCTCATGGAGAAACACTACTTGCAGAAACCTT 1081
Db 911 ---TAAAGAAAGAGGGGTATCAGTCTAATCTCATGGAGAAACACTACTTGCAGAAACCTT 967
Qy 1082 CTTAAGAGATGCTCTTTTATTGCTTACAAATGATTTCTAGTCTTTTAAACAACTGTTGAG 1141
Db 968 CTTAAGAGATGCTCTTTTATTGCTTACAAATGATTTCTAGTCTTTTAAACAACTGTTGAG 1027
Qy 1142 ATTTGTTTTAGTTGGTCGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201
Db 1028 ATTTGTTTTAGTTGGTCGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
Qy 1202 TTACCACTACTACATGCTGCGCAAGGTAAGGATCAGAGGACTGAAACAAATGATTTCTGCAA 1261
Db 1088 TTACCACTACTACATGCTGCGCAAGGTAAGGATCAGAGGACTGAAACAAATGATTTCTGCAA 1147
Qy 1262 CTCCTTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCCCTTAATAAATGTC 1317
Db 1148 CTCCTTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCCCTTAATAAATGTC 1203

RESULT 5

US-09-904-615-17
; Sequence 17, Application US/09904615
; Patent No. US2002026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT FILING DATE: 2001-07-16
; CURRENT FILING DATE: 2001-07-16
; PRIOR FILING DATE: 2000-02-23

```

; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1279)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1286)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-17

Query Match      82.3%; Score 1084.4; DB 10; Length 1291;
Best Local Similarity 97.8%; Pred. No. 6.2e-297;
Matches 1111; Conservative 1; Mismatches 17; Indels 7; Gaps 1;

QY 182 AAGAAAGAAACAAACAGAGAGGAAATGAAAAAGACATAATGATGTCATCCCAAGCCAAACA 241
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 75 AAAGAGGAACATAGAGGTGCCAAGGAACAAAGAACATAATGATGTCATCCCAAGCCAAACA 134

QY 242 AGCCATGCTGAAGTAAATGAACCATACCCACCCTTACCCACCAAGCAGCTTTATGGCT 301
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 135 AGCCATGCTGAAGTAAATGAACCATACCCACCCTTACCCACCAAGCAGCTTTATGGCT 194

QY 302 CCTGATTTCAACACCCCTCTGGGTTCAATCACTTAGAAAACCAAGCTCAGGGTCTCAG 361
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 195 CTTGATTTCAACAGCCCTCTGGGTTCAATCACTTAGAAAACCAAGCTCAGGGTCTCAG 254

QY 362 CTGCTCAGCCCTACGGCATCATCTCCGGGAATCTTGTAGCAGTCAACCGGGTCAA 421
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 255 CTGCTCAGCCCTACGGCATCATCTCCGGGAATCTTGTAGCAGTCAACCGGGTCAA 314

QY 422 GGAATATACAATGATAAATCCAAGTGTGGGAACAGCAGTAAATGAACCTTTAAAGAA 481
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 315 GGAATATACAATGATAAATCCAAGTGTGGGAACAGCAGTAAATGAACCTTTAAAGAA 374

QY 482 GCAAGGCACTAGGGGTGATCCAGATCATGTTGGATTCATGCATCATGGTTTGAAT 541
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 375 GCAAGGCACTAGGGGTGATCCAGATCATGTTGGATTCATGCATCATGGTTTGAAT 434

QY 542 GTTTTGTGTTAATATCTCTCTTTTAGAAGATATTAGGTTTTCCTCTACTGCTGT 601
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 435 GTTTTGTGTTAATATCTCTCTTTTAGAAGATATTAGGTTTTCCTCTACTGCTGT 494

QY 602 ATTGTGGATACCCATCTGGGGTGGCCCTTTCTTTTATATCTCTGGCTCTCTCTGTG 661
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 495 ATTGTGGATACCCATCTGGGGTGGCCCTTTCTTTTATATCTCTGGCTCTCTCTGTG 554

QY 662 TCAGCATCAAGGAGCTTCCCGTTGCTGTGGTGAAGCAGCCTGGGAATGAACATTGT 721
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 555 TCAGCATCAAGGAGCTTCCCGTTGCTGTGGTGAAGCAGCCTGGGAATGAACATTGT 614

QY 722 AGTTCTATCTTGGCCCTTCATTCGGAGTGTCTGCTGGTGGATATGTGCATCAATGGG 781
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 615 AGTTCTATCTTGGCCCTTCATTCGGAGTGTCTGCTGGTGGATATGTGCATCAATGGG 674

QY 782 GTAGTGGCCAAAGACTACTGGCCCGTCTTTCTGGAAGAGGATTTCAAGCCACGCTGATG 841
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 675 GTAGTGGCCAAAGACTACTGGCCCGTCTTTCTGGAAGAGGATTTCAAGCCACGCTGATG 734

QY 842 ATCTTCTCCCTCTTGGAGTCTTTCGTAGCTTGTGCCACAGCCCATTTTGCACCAAGCA 901
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 735 ATCTTCTCCCTCTTGGAGTCTTTCGTAGCTTGTGCCACAGCCCATTTTGCACCAAGCA 794
```

```

QY 902 AACACCACAACCAATATGTCTCTGGTTATTCCAAATATATGATGAAGCAACCTGTG 961
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 795 AACACCACAACCAATATGTCTCTGGTTATTCCAAATATATGATGAAGCAACCTGTG 854
QY 962 ACACGAGCGTCTTTCAGCTCCTCCAGATGCAACAACTACTCAGCTAATGCCCTAAA 1021
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 855 ACACGAGCGTCTTTCAGCTCCTCCAGATGCAACAACTACTCAGCTAATGCCCTAAA 910
QY 1022 TAGTAAAGAAAAGGGGTATCAGTCTATCTATCTGAGAAAACACTACTTGCAAAACCT 1081
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 911 ---TAAAGAAAAGGGGTATCAGTCTATCTATCTGAGAAAACACTACTTGCAAAACCT 967
QY 1082 CTTAAGAAAGATGCTCTTTTATCTGTACAAATGATTTCTAGTCTTTAAAACTGTGTTGAG 1141
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 968 CTTAAGAAAGATGCTCTTTTATCTGTCTACAAATGATTTCTAGTCTTTAAAACTGTGTTGAG 1027
QY 1142 ATTTGTTTTAGTGGTGGCTAAATGATGGCTGTATCTCCCTTCACTGTCTCTTCTCTACA 1201
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1028 ATTTGTTTTAGTGGTGGCTAAATGATGGCTGTATCTCCCTTCACTGTCTCTTCTCTACA 1087
QY 1202 TTACCACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAAATGATCTTGCAA 1261
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1088 TTACCACTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAAATGATCTTGCAA 1147
QY 1262 CTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCTCTTAATAAATGTC 1317
DB 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1148 CTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTACTTTTCTCTTAATAAATGTC 1203

RESULT 6
US-10-054-988-61
; Sequence 61, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/10/054,988
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1280)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1291)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-054-988-61
```

Query Match 81.9%; Score 1079; DB 9; Length 1292;

Best Local Similarity 97.3%; Pred. No. 2.le-295; Matches 1105; Conservative 4; Mismatches 20; Indels 7; Gaps 1;									
Qy	182	AAGAAAGAAACAAACAGAGAGAGAAATGAAAGACATAATGATGTCATCCAAAGCCAACA	241						
Db	76	ANAGAGGAACATAGAGTGCCAAAGAAAGAAACAAACATAATGATGTCATCCAAAGCCAACA	135						
Qy	242	AGCCATGCTGAAGTAAATGAAACCATATACCCAAACCTTACCACCAAGCAGCTTTATGGCT	301						
Db	136	AGCCATGCTGAAGTAAATGAAACCATATACCCAAACCTTACCACCAAGCAGCTTTATGGCT	195						
Qy	302	CCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTGCTCAG	361						
Db	196	CCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTGCTCAG	255						
Qy	362	CGTGCTCAGCCCTACGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCGGGTCAA	421						
Db	256	CGTGCTCAGCCCTACGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCGGGTCAA	315						
Qy	422	GGAATATACAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTTAAAGAGAA	481						
Db	316	GGAATATACAATGATAATCCAAAGTGTGGGAACAGCAGTAATGAACCTTTTAAAGAGAA	375						
Qy	482	GCAAGGCACCTAGGGGTATCCAGATCATGTTGGATTGATGACATTTGGTAATT	541						
Db	376	GCAAGGCACCTAGGGGTATCCAGATCATGTTGGATTGATGACATTTGGTAATT	435						
Qy	542	GTTTGTGTTTAAATATCTCTCTTTAGACAAGTATTAGGTTTGGCTCTACTGCTGTT	601						
Db	436	GTTTGTGTTTAAATATCTCTCTTTAGACAAGTATTAGGTTTGGCTCTACTGCTGKT	495						
Qy	602	ATTGGTGGATACCCATTCTGGGGTGGCTTTCTTTTATATCTCTGGCTCTCTCTGTG	661						
Db	496	ATTGGTGGATACCCATTCTGGGGTGGCTTTCTTTTATATCTCTGGCTCTCTCTGTG	555						
Qy	662	TCAGATCCAGAGAGCTTTCCCGTTGCTGTGTGAAGCAGCCCTGGGAATGAACATGTT	721						
Db	556	TCAGATCCAGAGAGCTTTCCCGTTGCTGTGTGAAGCAGCCCTGGGAATGAACATGTT	615						
Qy	722	AGTTCTATCTTGGCTTCATTGGAGTGATCTGCTGCTGGTGATATGTGTCATCAATGGG	781						
Db	616	AGGTCTATCTTGGCTTCATTGGAGTGATCTGCTGCTGGTGATATGTGTCATCAATGGG	675						
Qy	782	GTAGCTGCCAAGACTACTGGCCGCTGCTTTCTGAAAAGGCATTTTCAGCCAGCTGATG	841						
Db	676	GTARCTGCCAAGACTACTGNCCTGCTTTCTGAAAAGGCATTTTCAGCCAGCTGATG	735						
Qy	842	ATCTTCTCCCTCTTGGAGTCTTCTGATAGCTTGTGCCACAGCCCATTTGCCAACCAAGCA	901						
Db	736	ATCTTCTCCCTCTTGGAGTCTTCTGATAGCTTGTGCCACAGCCCATTTGCCAACCAAGCA	795						
Qy	902	AACACCACAACCAATATGCTGCTGCTGTTTATCCAAATATGATGAAAGCAACCTGTG	961						
Db	796	AACACCACAACCAATATGCTGCTGCTGTTTATCCAAATATGATGAAAGCAACCTGTG	855						
Qy	962	ACACCAGGCTCTTCTCAGCTCCTCCAGATGCAACAACTACTCAGCTAATGCCCTAAA	1021						
Db	856	ACACCAGGCTCTTCTCAGCTCCTCCAGATGCAACAACTACTCAGCTAATGCCCTAAA	911						
Qy	1022	TAGTAAAGAAAGAGGGGTATCAGTCTAATCTCATGTGGAGAAAACCTACTTGCAAAACCTT	1081						
Db	912	---TAAAGAAAGAGGGGTATCAGTCTAATCTCATGTGGAGAAAACCTACTTGCAAAACCTT	968						
Qy	1082	CTTAAGAGAGATGCTTTTATTGTTCTACAAATGATTTCTAGTCTTTTAAAACTGTCTTGAG	1141						
Db	969	CTTAAGAGAGATGCTTTTATTGTTCTACAAATGATTTCTAGTCTTTTAAAACTGTCTTGAG	1028						
Qy	1142	ATTTGTTTTTATGTTGGTCGTAATGATGCTGATATCTCCCTTCACTGTCTCTTCTTACA	1201						
Db	1029	ATTTGTTTTTATGTTGGTCGTAATGATGCTGATATCTCCCTTCACTGTCTCTTCTTACA	1088						
Qy	1202	TTACCACCTACTACATGCTGGCAAGGTCGAGGATCAGAGGACTGAAAATGATTTCTGCAA	1261						

Db	1089	TTACCACCTACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAATGATTTCTGCAA	1148						
Qy	1262	CTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTACTTTTCCCTTAATAAAATGTC	1317						
Db	1149	CTCTCTTAAAGTTAGAAATGTTCTGTTTCATATTACTTTTCCCTTAATAAAATGTC	1204						
RESULT 7									
US-09-739-254-61									
; Sequence 61, Application US/09739254									
; Patent No. US20010021700A1									
; GENERAL INFORMATION:									
; APPLICANT: Rosen et al.									
; TITLE OF INVENTION: 49 Human Secreted Proteins									
; FILE REFERENCE: P2032P1									
; CURRENT APPLICATION NUMBER: US/09/739,254									
; CURRENT FILING DATE: 2000-12-19									
; EARLIER APPLICATION NUMBER: 09/511,554									
; EARLIER FILING DATE: 2000-02-23									
; EARLIER APPLICATION NUMBER: PCT/US99/19330									
; EARLIER FILING DATE: 1999-08-24									
; EARLIER APPLICATION NUMBER: 60/097,917									
; EARLIER FILING DATE: 1998-08-25									
; EARLIER APPLICATION NUMBER: 60/098,634									
; EARLIER FILING DATE: 1998-08-31									
; NUMBER OF SEQ ID NOS: 170									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 61									
; LENGTH: 1292									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: SITE									
; LOCATION: (71)									
; OTHER INFORMATION: n equals a,t,g, or c									
; FEATURE:									
; NAME/KEY: SITE									
; LOCATION: (697)									
; OTHER INFORMATION: n equals a,t,g, or c									
; FEATURE:									
; NAME/KEY: SITE									
; LOCATION: (1280)									
; OTHER INFORMATION: n equals a,t,g, or c									
; FEATURE:									
; NAME/KEY: SITE									
; LOCATION: (1287)									
; OTHER INFORMATION: n equals a,t,g, or c									
; FEATURE:									
; NAME/KEY: SITE									
; LOCATION: (1291)									
; OTHER INFORMATION: n equals a,t,g, or c									
US-09-739-254-61									
Query Match 81.9%; Score 1079; DB 10; Length 1292;									
Best Local Similarity 97.3%; Pred. No. 2.le-295;									
Matches 1105; Conservative 4; Mismatches 20; Indels 7; Gaps 1;									
Qy	182	AAGAAAGAAACAAACAGAGAGGAATGAAAAGACATAATGATGTCTATCCAAAGCCAACA	241						
Db	76	AAAGAGGAACATAGAGGTGCCAAAGGAAAGGAACAAAGACATAATGATGTCTATCCAAAGCCAACA	135						
Qy	242	AGCCATGCTGAAGTAAATGAAACCATATCCCAACCCCTTACCACCAAGCAGCTTTATGGCT	301						
Db	136	AGCCATGCTGAAGTAAATGAAACCATATCCCAACCCCTTACCACCAAGCAGCTTTATGGCT	195						
Qy	302	CCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTGCTCAG	361						
Db	196	CCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAACCAAGCTCAGGGTGCTCAG	255						
Qy	362	CGTGCTCAGCCCTACGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCGGGTCAA	421						
Db	256	CGTGCTCAGCCCTACGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCAACCGGGTCAA	315						

Db 676 GTARCTGCCAAGACTACTGGNCCGTCTTTCTGGAAAAGGCATTTTCAGCCACGCTGATG 735
QY 842 ATCTTCTCCCTCTCTTGGAGTCTTTCGTAGCTTGTGCACAGCCCATTTTGGCAACCAAGCA 901
Db 736 ATCTTCTCCYCTTGGAGTCTTTCGTAGCTTGTGCACAGCCCATTTTGGCAACCAAGCA 795
QY 902 AACACCACAACCAATATGCTGTCTGCTGTTTATTCCAAATATGTTATGAAGCAACCTGTG 961
Db 796 AACACCACAACCAATATGCTGTCTGCTGTTTATTCCAAATATGTTATGAAGCAACCTGTG 855
QY 962 ACACCAGGCTCTTTCAGCTCCTCCAGAGTCAACAACACTACTAGCTAATGCCCCCTAAA 1021
Db 856 ACACCAGGCTCTTTCAGCTCCTCCAGAGTCAACAACACTACTAGCTAATGCCCCCTAAA 911
QY 1022 TAGTAAAGAAAAGGGGTATCAGTCAATCTCATGAGGAGAAAACACTTTCGCAAACTT 1081
Db 912 ---TAAAGAAAAGGGGTATCAGTCAATCTCATGAGGAGAAAACACTTTCGCAAACTT 968
QY 1082 CTTAAGAAGATGCTCTTTTATTGTCTACAAATGATTTCTAGTCTTTTAAAACTGTGTTGAG 1141
Db 969 CTTAAGAAGATGCTCTTTTATTGTCTACAAATGATTTCTAGTCTTTTAAAACTGTGTTGAG 1028
QY 1142 ATTTGTTTTAGTTGGTGGCTAATGATGGGTGATCTCCCTTCACTGCTCTCTCTCTACA 1201
Db 1029 ATTTGTTTTAGTTGGTGGCTAATGATGGGTGATCTCCCTTCACTGCTCTCTCTCTACA 1088
QY 1202 TTACCACACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAATGATTTCTGCAA 1261
Db 1089 TTACCACACTACATGCTGGCAAGGTGAAGGATCAGAGGACTGAAAATGATTTCTGCAA 1148
QY 1262 CTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTTCTTTTCCCTTAAATAAATGTC 1317
Db 1149 CTCTCTTAAAGTTAGAAATGTTTCTGTTTCATATTTCTTTTCCCTTAAATAAATGTC 1204

RESULT 9

US-09-925-299-727
; Sequence 727, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 727
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (321)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (405)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (422)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (442)

; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-727
Query Match 29.6%; Score 389.6; DB 9; Length 441;
Best Local Similarity 96.6%; Pred. No. 3.9e-100;
Matches 427; Conservative 0; Mismatches 10; Indels 5; Gaps 3;
QY 113 GAAGACAAGTCTTCTGACTCTATGTTGAGGCCAGTTTGAATAATGAGGGAGAATAAACCAT 172
Db 2 GAAGACAAGTCTTCTGACTCTATGTTGAGGCCAGTTTGAATAATGAGGGAGAATAAACCAT 61
QY 173 GAAGCAAAACAAGAAAGAAACAACAGAGAGGAATGAAAAGACATAATGATGTCATCC 232
Db 62 GAAGCAAAACAAGAAAGAAACAACAGAGAGGAATGAAAAGACATAATGATGTCATCC 121
QY 233 AAGCCAACAAGCCATGCTGAAGTAATGAACCATACCAACCCCTTACCCACCAAGCAGC 292
Db 122 AAGCCAACAAGCCATGCTGAAGTAATGAACCATACCAACCCCTTACCCACCAAGCAGC 181
QY 293 TTTATGGCTCTCTGGATTTCAACAGCCTCTGGGTTTCAATCAACTTAGAAAACCAAGCTCAG 352
Db 182 TTTATGGCTCTCTGGATTTCAACAGCCTCTGGGTTTCAATCAACTTAGAAAACCAAGCTCAG 241
QY 353 GGTGCTACGCTGCTCAGCCCTACGCATCATCTCCGGGAATCTTTGCTAGCAGTCAA 412
Db 242 GGTGCTACGCTGCTCAGCCCTACGCATCATCTCCGGGAATCTTTGCTAGCAGTCAA 301
QY 413 CCGGTCAAGGAATATACAAATGATAATCCAAGTGTGGGAACAGCAGTGAACATTT 472
Db 302 CCGGTCAAGGAATATACAAATGATAATCCAAGTGTGGGAACAGCAGTGAACATTT 361
QY 473 AAA--GAAGAGCAAGGCACCTA--CGGTGTATCCAGATCATGTTGGATTGATGCACATTT 529
Db 362 AAAAGAAAGCAAGGCACCTA--CGGTGTATCCAGATCATGTTGGATTGATGCACATTT 419
QY 530 GGTGTTGGAATGTTTGTGTT 551
Db 420 TGNTTGGAAATGNTTGTGTT 441

RESULT 10

US-09-925-299-727
; Sequence 727, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 727
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (321)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (405)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (422)

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-727

```

Query Match	29.5%	Score 389.6	DB 10	Length 441
Best Local Similarity	96.6%	Pred. No. 3.9e-100		
Matches 427	Conservative 0	Mismatches 10	Indels 5	Gaps 3
QY	113	GAAGACAAGTCTCTTGACTCTATGTTGAGGCCAGTTGAAAAATGAGGAGAGAAATAAACCAT	172	
DB	2	GAAGACAAGTCTCTTGACTCTATGTTGAGGCCAGTTGAAAAATGAGGAGAGAAATAAACCAT	61	
QY	173	GAACGAAACAAGAAAGAAACAAACAGAGAGAGGAATGAAAAGACATAATGATGTCATCC	233	
DB	62	GAACGAAACAAGAAAGAAACAAACAGAGAGAGGAATGAAAAGACATAATGATGTCATCC	121	
QY	233	AAGCCAAACAGCCATGCTGAAAGTAAATGAAACCATATCCCAACCCCTTACCACCAACGAGCG	292	
DB	122	AAGCCAAACAGCCATGCTGAAAGTAAATGAAACCATATCCCAACCCCTTACCACCAACGAGCG	181	
QY	293	TTTATGGCTCTCGATTTTCAACAGCCTCTCGGGTTCAATCAACTTAGAAAAACCAAGCTCAG	352	
DB	182	TTTATGGCTCTCGATTTTCAACAGCCTCTCGGGTTCAATCAACTTAGAAAAACCAAGCTCAG	241	
QY	353	GGTGCTCAGCGTGCTCAGCCCTCAGGCATCACATCTCCGGGAATCTTTGCTAGACAGTCAA	412	
DB	242	GGTGCTCAGCGTGCTCAGCCCTCAGGCATCACATCTCCGGGAATCTTTGCTAGACAGTCAA	301	
QY	413	CCGGGTCAAGGAATAATACAAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACATTT	472	
DB	302	CCGGGTCAAGGAATAATACAAATGATAAATCCAAAGTGTGGGAACAGCAGTAATGAACATTT	361	
QY	473	AAA - GAAGAAGCAAAAGGCACATA - GGGGTGATCCAGATCATGTTGGATTCATGCACATTT	529	
DB	362	AAAAGAAAGCAAGGCACATAAGGGGTGATNCAGATCATGTTGGATTCATGCACATTT - CCAT	419	
QY	530	GGTTTTGGAATTTGTTTGCTT	551	
DB	420	TGNTTTTGGAAATGNTTNGCTT	441	

```

RESULT 11
US - Sequence 494, Application US/09803719
: Sequence 494, Application US/09803719
: Publication No. US20030044783A1
: GENERAL INFORMATION:
: APPLICANT: Williams, Lewis T.
: APPLICANT: Escobedo, Jaime
: APPLICANT: Innis, Michael A.
: APPLICANT: Garcia, Pablo Dominiguez
: APPLICANT: Sudduth-Klinger, Julie
: APPLICANT: Reinhard, Christoph
: APPLICANT: Glese, Klaus
: APPLICANT: Randazzo, Filippo
: APPLICANT: Kennedy, Giulia C.
: APPLICANT: Pot, David
: APPLICANT: Kassam, Altaf
: APPLICANT: Lamson, George
: APPLICANT: Drmanac, Radoje
: APPLICANT: Crkvenjakov, Radomir
: APPLICANT: Dickson, Mark
: APPLICANT: Drmanac, Snezana
: APPLICANT: Labat, Ivan
: APPLICANT: Leshkowitz, Dena
: APPLICANT: Kita, David
: APPLICANT: Garcia, Veronica
: APPLICANT: Jones, Lee William
: APPLICANT: Stache-Grain, Birgit

```

```

; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-494

```

Query Match	23.1%	Score 303.6	DB 9	Length 351
Best Local Similarity	93.0%	Pred. No. 8.5e-76		
Matches 318	Conservative 0	Mismatches 24	Indels 0	Gaps 0
Qy	172	TGACGAAACAGAAAGAAACAAACAAACAGAGAGGGAATGAAAAGACATAAATGATGTCATC	231	
Db	10	TGGAGCAGAAAGAGGAAACATAGAGGTGCCAAGGAACAAAGACATAATGATGTCATC	69	
Qy	232	CAAGCCAAACAGCCATGCTGAAGTAAATGAAACCATACCCAAACCTTACCCACCAAGCAG	291	
Db	70	CAAGCCAAACAGCCATGCTGAGTAAATGAAACCATACCCAAACCTTACCCACCAAGCAG	129	
Qy	292	CTTTATGGCTCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAAACCAAGTC	351	
Db	130	CTTTATGGCTCTGGATTTCAACAGCCTCTGGGTTCAATCAACTTAGAAAAACCAAGTC	189	
Qy	352	GGGTGCTCAGCGTGCTCAGCCCTACGCGATCACATCTCCGGGAATCTTTGCTAGCAGTCA	411	
Db	190	GGGTGCTCAGCGTGCTCAGCCCTATGGCATCACATCTCCGGGAATCTTTGCTAGCAGTCA	249	
Qy	412	ACCGGGTCAAGGAAATATACAAATGATAAATCCAAGTGTGGGAACAGCAGTAATGAACCTT	471	
Db	250	ACCGGGTCAAGGAAATATATAAATGATAAATCCAAGTGTGGGAACAGCAGTAATGAACCTT	309	
Qy	472	TAAAGAGAAGCAAGGCATAGGGGTGATCCAGATCATGGT	513	
Db	310	TAAAGAGAAGCAAGGCATAGGGGTGATCCAGATCATGGT	351	

RESULT 12

US-09-803-719-638

US-09-803-719-638, Application US/09803719

Sequence 638, Application US/09803719

Publication No. US2003004783A1

GENERAL INFORMATION:

APPLICANT: Williams, Lewis T.

APPLICANT: Escobedo, Jaime

APPLICANT: Innis, Michael A.

APPLICANT: Garcia, Pablo Dominguez

APPLICANT: Sudduth-Klinger, Julie

APPLICANT: Reinhard, Christoph

APPLICANT: Giese, Klaus

APPLICANT: Randazzo, Filippo

APPLICANT: Kennedy, Giulia C.

APPLICANT: Pot, David

APPLICANT: Kassam, Altaf

APPLICANT: Lamson, George

APPLICANT: Drmanac, Radoje

APPLICANT: Cirkvenjakov, Radomir

APPLICANT: Dickson, Mark

APPLICANT: Drmanac, Snezana

APPLICANT: Labat, Ivan

APPLICANT: Leshkowitz, Pena

APPLICANT: Kita, David

APPLICANT: Garcia, Veronica

APPLICANT: Jones, Lee William

APPLICANT: Stache-Crain, Birgit

TITLE OF INVENTION: Human Genes and Gene Products

FILE REFERENCE: 1624.002

CURRENT APPLICATION NUMBER: US/09/803.719

Query Match 7.0%; Score 92.6; DB 10; Length 125;

Best Local Similarity 87.2%; Pred. No. 3.1e-16;
Matches 109; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

```
QY 508 CATGGTTGGATTGATGCACATGGGT-TTTGGAAATTCCTTTGGTAAATATCCTTCTCTT 566
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 CGTGGTTGGATTGATGCACATGGGTCTCGGNAACGACCTGNGTNNAATATCCTTCTCTT 60
QY 567 TTAGAGAAGTATTAGGTGTTTGGCTCTACTGCTGTTATTGGTGATACCCATTCCTGGGGTG 626
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 TTAGAGAAGTATTAGGTGTTTGGCTCTACTGCTGTTATTGGTGATACCCATTCCTGGGGG 120
QY 627 GCCTT 631
Db | | | |
121 GCCTT 125
```

RESULT 15

```
US-10-017-754-2003
; Sequence 2003, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 21021.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2003
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-2003
```

Query Match 6.2%; Score 81.2; DB 9; Length 753;

Best Local Similarity 58.3%; Pred. No. 1.8e-12;
Matches 190; Conservative 0; Mismatches 118; Indels 18; Gaps 2;

```
QY 590 TCTACTGCTGTTATTGGTGATACCCATTCCTGGGTGGCTTTCTTTTATATCTCTGCG 649
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
298 TCTATTTCATCTACGAGGCTTCCCTCTCTGGGAGGCTTGGGTTTATCATTTTCAGGA 357
QY 650 TCTCTCTCTGTGTGACATCCAGAGAGCT---TTCCCGTTGTCTGCTGAAAGGACGCTG 706
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
358 TCTCTCTCCGTGGCAGCAGAAAATCAGCCATATCTTATTGCTGCTGTCTGCGAGTTG 417
QY 707 GGAATCAACATGTTAGTTCTATCTTGGCCCTTCATTTGGAGTGATTCCTGCTGGTGGAT 766
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
418 GGCTTGAACATCGTCAGTCAATCTGCTCTGCAGTTGGAGTCATCTCTTCATCACAGAT 477
QY 767 ATGTGCAT-----CAATGGGTAGCTGGCCAGCACTACTGGGCGGTGCTT 811
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
478 CTAAGTATTCCTCCACCCCATATGCTACCCGAGCTATTATCTTACCGCTGGGTGTGAAC 537
QY 812 TCTGGAAGGCAATTTTCAGCCAGCTGATGATCTCTCCCTCTTGGAGTTCTTCGTAGCT 871
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
538 CTTGGAATGGCGATTTCTGGCGTCTGCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTG 597
QY 872 TGTGCCACAGCCCATTTTGGCCACCA 897
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
598 TGGCATCTTCCCACTTTGGCTGCCA 623
```

Search completed: July 16, 2003, 18:04:17
Job time : 176.297 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: July 16, 2003, 03:49:46 ; Search time 1060.5 Seconds
(without alignments)
15998.962 Million cell updates/sec

Title: US-09-981-353-114
Perfect score: 583
Sequence: 1 cccgtgaggctgccttg.....tatgtgaatgaataagta 583

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pi.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_ron.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	304.4	52.2	465	9	BC029475	BC029475 Homo sapi
2	277.8	47.7	454	9	BC032338	BC032338 Homo sapi
3	277.4	47.6	415	6	AX330310	AX330310 Sequence
4	277.4	47.6	415	6	AX336406	AX336406 Sequence
5	277.4	47.6	415	6	AX411150	AX411150 Sequence
6	277.4	47.6	415	9	HSMT1L	X76717 H.sapiens M
7	276.2	47.4	404	9	BC018190	BC018190 Homo sapi
8	260.2	44.6	120825	9	HSJ733M16	AL109627 Human DNA
9	259.8	44.6	368	9	AY077636	AY077636 Homo sapi
10	248	42.5	72930	9	HS1174N9	AL031602 Human DNA
11	246.4	42.3	422	9	BC008408	BC008408 Homo sapi
12	245.2	42.1	409	6	AX330235	AX330235 Sequence
13	245.2	42.1	409	6	AX331811	AX331811 Sequence
14	245.2	42.1	409	6	AX336059	AX336059 Sequence
15	245.2	42.1	409	6	AX336280	AX336280 Sequence
16	245.2	42.1	409	6	AX336680	AX336680 Sequence
17	245.2	42.1	409	6	AX410056	AX410056 Sequence
18	237	40.7	844	9	HUMMET1PG	MI3073 Human metal
19	234.4	40.2	729	9	BC009699	BC009699 Homo sapi
20	233.2	40.0	288	11	G06180	G06180 human STS W
21	227.2	39.0	468	9	BC029453	BC029453 Homo sapi
22	226.8	38.9	1071	9	AF078844	AF078844 Homo sapi
23	226.6	38.9	164529	2	AC079975	AC079975 Homo sapi
24	226.6	38.9	164933	2	AC074231	AC074231 Homo sapi
25	226.6	38.9	172753	9	AL359259	AL359259 Human DNA
26	226.6	38.9	176671	2	AC080149	AC080149 Homo sapi
27	221.8	38.0	434	9	BC020757	BC020757 Homo sapi
28	221.8	38.0	453	6	AX399847	AX399847 Sequence
29	219.8	37.7	367	6	AX336464	AX336464 Sequence
30	219.8	37.7	367	6	AX411116	AX411116 Sequence
31	219.8	37.7	367	6	HSFNET	X84177 H.sapiens m
32	219.4	37.6	431	9	BC028280	BC028280 Homo sapi
33	219.4	37.6	160716	2	AC053475	AC053475 Homo sapi
34	219.4	37.6	176668	2	AC027096	AC027096 Homo sapi
35	219.4	37.6	189643	9	AL161729	AL161729 Human DNA
36	216	37.0	646	6	AX067316	AX067316 Sequence
37	216	37.0	738	9	AB064124	AB064124 Homo sapi
38	216	37.0	761	9	AB064049	AB064049 Homo sapi
39	216	37.0	781	9	AB064131	AB064131 Homo sapi
40	216	37.0	786	9	AB064060	AB064060 Homo sapi
41	216	37.0	791	9	AB064135	AB064135 Homo sapi
42	216	37.0	794	9	AB064122	AB064122 Homo sapi
43	216	37.0	799	9	AB064129	AB064129 Homo sapi
44	216	37.0	796	9	AB064120	AB064120 Homo sapi
45	216	37.0	796	9	AB064128	AB064128 Homo sapi

ALIGNMENTS

RESULT 1
BC029475

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BC029475 465 bp mRNA linear PRI 16-MAY-2002
Homo sapiens, Similar to RNA helicase-related protein, clone
MGC:32848 IMAGE:4723818, mRNA, complete cds.

BC029475

BC029475.1 GI:20809974

MGC.

Homo sapiens.

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 465)

Strausberg, R.

Direct Submission


```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3797 11-APR-2002;
GENE LOGIC INC (US)
FEATURES Location/Qualifiers
source 1. 415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. X76717"
BASE COUNT 110 a 102 c 87 g 115 t 1 others
ORIGIN
Query Match 47.6%; Score 277.4; DB 6; Length 415;
Best Local Similarity 88.3%; Pred. No. 9e-67;
Matches 326; Conservative 1; Mismatches 32; Indels 10; Gaps 2;
QY 221 CCGCTGCGTGTTCCTCTTGCATCGGGAACCTCCTGCTTCTCTTGCCTCGAAATGGAGCC 280
DB 1 CCGCTGCGTGTTCCTCTTGCATCGGGAACCTCCTGCTTCTCTTGCCTCGAAATGGAGCC 60
QY 281 CAATGCTCCTGCTGCGCTGTGGCTCCTGTGCTGCGCGCTCCTGCAAAATGCAAGA 340
DB 61 CAATGCTCCTGCTGCGCTGTGGCTCCTGTGCTGCGCGCTCCTGCAAAATGCAAGA 120
QY 341 GTGCAATGCACTCCTCGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
DB 121 GTGCAATGCACTCCTCGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 401 GTGTCGCCAGGCTGCATCTGCAAGGGGCATCAGAGAGTGCAGTGCCTGCTGCTGATG 460
DB 181 GTGTCGCCAGGCTGCATCTGCAAGGGGCATCAGAGAGTGCAGTGCCTGCTGCTGATG 240
QY 461 TCCGGACAGCCTGCTCGAAGATATAGAAAGAGTGACCTGCACAACTTGA----- 512
DB 241 CCAGGACAGCTGTGCTCTCAGATGTAATAGAGCAACCTATATAAACCTGGATTTTTT 300
QY 513 ATTTTTCATACAAACCTGCCCCCATCTACT--GTATTTTTTTTAAATGAATATGGA 570
DB 301 TTTTTCATACAAACCTGCCCCCATCTACT--GTATTTTTTTTAAATGAATATGGA 360
QY 571 ATGATAATA 579
DB 361 ATGGCAATA 369
RESULT 6
HSM11L HSM11L 415 bp mRNA linear PRI 10-JAN-1994
LOCUS H.sapiens MT-11 mRNA.
DEFINITION H.sapiens MT-11 mRNA.
ACCESSION X76717
VERSION X76717.1 GI:435674
KEYWORDS MT-11 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 415)
AUTHORS Soumilleon, A., Van Weyenbergh, J. and De Ley, M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 415)
AUTHORS De Ley, M.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-1993) M. De Ley, Katholieke Universiteit Leuven,
Lab voor Biochemie, Dekenstraat 6, 3000 Leuven, BELGIUM
FEATURES Location/Qualifiers
source 1. 415
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pBlue-Mt-11"
```

```
/tissue_type="liver"
/clone_lib="lambdazap-MT-11"
/dev_stage="fetus"
53...238
/codon_start=1
/product="MT-11 protein"
/protein_id="CAA54136.1"
/db_xref="GI:435675"
/db_xref="SWISS-PROT:P80297"
/translation="MDPNCSPVSCACAGSCKECKTSCCKSCSCCPVCACAK
ADGCKGTSKSCCA"
polya_signal 366...371
BASE COUNT 110 a 102 c 87 g 115 t 1 others
ORIGIN
Query Match 47.6%; Score 277.4; DB 9; Length 415;
Best Local Similarity 88.3%; Pred. No. 9e-67;
Matches 326; Conservative 1; Mismatches 32; Indels 10; Gaps 2;
QY 221 CCGCTGCGTGTTCCTCTTGCATCGGGAACCTCCTGCTTCTCTTGCCTCGAAATGGAGCC 280
DB 1 CCGCTGCGTGTTCCTCTTGCATCGGGAACCTCCTGCTTCTCTTGCCTCGAAATGGAGCC 60
QY 281 CAATGCTCCTGCTGCGCTGTGGCTCCTGTGCTGCGCGCTCCTGCAAAATGCAAGA 340
DB 61 CAATGCTCCTGCTGCGCTGTGGCTCCTGTGCTGCGCGCTCCTGCAAAATGCAAGA 120
QY 341 GTGCAATGCACTCCTCGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
DB 121 GTGCAATGCACTCCTCGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 401 GTGTCGCCAGGCTGCATCTGCAAGGGGCATCAGAGAGTGCAGTGCCTGCTGCTGATG 460
DB 181 GTGTCGCCAGGCTGCATCTGCAAGGGGCATCAGAGAGTGCAGTGCCTGCTGCTGATG 240
QY 461 TCCGGACAGCCTGCTCGAAGATATAGAAAGAGTGACCTGCACAACTTGA----- 512
DB 241 CCAGGACAGCTGTGCTCTCAGATGTAATAGAGCAACCTATATAAACCTGGATTTTTT 300
QY 513 ATTTTTCATACAAACCTGCCCCCATCTACT--GTATTTTTTTTAAATGAATATGGA 570
DB 301 TTTTTCATACAAACCTGCCCCCATCTACT--GTATTTTTTTTAAATGAATATGGA 360
QY 571 ATGATAATA 579
DB 361 ATGGCAATA 369
RESULT 7
BC018190 404 bp mRNA linear PRI 06-DEC-2001
LOCUS Homo sapiens, Similar to metallothionein 1L, clone MGC:9187
DEFINITION IMAGE:3859643, mRNA, complete cds.
ACCESSION BC018190
VERSION BC018190.1 GI:17390422
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 404)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
TITLE Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobobcm.tmc.edu.
 Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 21 Row: d Column: 7
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505270.

FEATURES

Source

1. 404
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:9187 IMAGE:3859643"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH_MGC_66"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 57. 242
 /codon_start=1
 /product="Similar to metallothionein 1L"
 /protein_id="AAH18190.1"
 /db_xref="GI:17390423"
 /translation="MDPNCSPVSCAGSKCKNEKCTCKKSCCCPVGCAK
 AGCKIGKGSFSDKSCCA"

CDS

94 a 105 c 88 g. 117 t
 47.4%; Score 276.2; DB 9; Length 404;
 Best Local Similarity 88.3%; Pred. No. 1.9e-66;
 Matches 326; Conservative 0; Mismatches 33; Indels 10; Gaps 2;

BASE COUNT ORIGIN

Query Match 47.4%; Score 276.2; DB 9; Length 404;
 Best Local Similarity 88.3%; Pred. No. 1.9e-66;
 Matches 326; Conservative 0; Mismatches 33; Indels 10; Gaps 2;
 221 CCGCTGCGTGTTCCTCTTGATCGGGAACCTCTGCTCTCTCTGCTCGAATGACCC 280
 |||||
 5 CCGCTGCGTGTTCCTCTTGATCGGGAACCTCTGCTCTCTCTGCTCGAATGACCC 64
 |||||
 281 CAACTGCTCTGCTCGCCCTGTGGCTGTGCTGTGCTGTGCGGCTCTGCAATGCAAGA 340
 |||||
 65 CAACTGCTCTGCTCGCCCTGTGGCTGTGCTGTGCTGTGCGGCTCTGCAATGCAAGA 124
 |||||
 341 GTGCAATGCACCTCTCTGCAAGAGAGCTGCTCTCTGCTGCGGCTGTGCGGCTGTC 400
 |||||
 125 GTGCAATGCACCTCTCTGCAAGAGAGCTGCTCTCTGCTGCGGCTGTGCGGCTGTC 184
 |||||
 401 GTGTCGCCAGGCTGTCATCGAAAGGCGATCAGAGAGTGCAGCTGCTGTCCTGATG 460
 |||||
 185 GTGTCGCCAGGCTGTCATCGAAAGGCGATCAGAGAGTGCAGCTGCTGTCCTGATG 244
 |||||
 461 TCGGGACACCCCTGCTCGAAGATATAGAAGAGTGCACCTGCGACAACTTGA----- 512
 |||||
 245 CCAGGACAGCTGTGCTCTCAGATGTAAATAGACCACTATATAAACCTGGATTTT 304
 |||||
 513 ATTTTTCCTACACACCTGCCCATCTACT--CTATTTTTAAATGAATATGTGA 570
 |||||
 305 TTTTTCCTACACACCTGCCCATCTACT--CTATTTTTAAATGAATATGTGA 570
 |||||
 571 ATGATAATA 579
 |||||
 365 ATGGCAATA 373

RESULT 8

HSJ733M16

LOCUS

DEFINITION

Human DNA sequence from clone RP4-733M16 on chromosome 1p36.11-36.23 Contains a pseudogene similar to metallothionein 1L, an EST, STSs, GSSs and CpG islands, complete sequence.
 AL109627

ACCESSION

ALL109627.18 GI:6742145
 HTG: CpG Island; metallothionein.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 120825)
 Hall, R.
 Direct Submission
 Submitted (05-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 Requests: clonequest@sanger.ac.uk
 On Jan 24, 2000 this sequence version replaced gi:6723665.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence is the entire insert of clone RP4-733M16 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>
 RP4-733M16 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pCVPAC2.
 Location/Qualifiers
 1..120825
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p36.11-36.23"
 /clone="RP4-733M16"
 /clone_lib="RPCI-4"
 /clone_1lib="RPCI-4"
 /notes="MER75 repeat: matches 1. 208 of consensus"
 repeat_region
 238..444
 /notes="MER75 repeat: matches 1. 208 of consensus"
 repeat_region
 2440..2713
 /notes="AluSg1 repeat: matches 1. 275 of consensus"
 repeat_region
 3522..3597
 /notes="L1MC1 repeat: matches 6256. 6331 of consensus"
 repeat_region
 3902..4310
 /notes="L1MC1 repeat: matches 5848. 6256 of consensus"
 repeat_region
 4613..5146
 /notes="L1MC1 repeat: matches 5313. 5848 of consensus"
 repeat_region
 5147..5184
 /notes="19 copies 2 mer ta 94% conserved"
 repeat_region
 5845..5895
 /notes="L1MC1 repeat: matches 5276. 5330 of consensus"
 repeat_region
 5977..6150
 /notes="L1MC1 repeat: matches 5127. 5276 of consensus"
 repeat_region
 7076..7297
 /notes="L1MC1 repeat: matches 4890. 5132 of consensus"
 repeat_region
 7313..7633
 /notes="AluYa8 repeat: matches 1. 310 of consensus"
 repeat_region
 8356..8408
 /notes="L1MC1 repeat: matches 4850. 4900 of consensus"
 repeat_region
 8538..9097
 /notes="L1MC1 repeat: matches 4264. 4850 of consensus"
 misc_feature
 complement(9794..10350)
 /note="match: GSS: Em:AQ535906"
 repeat_region
 9824..9994

QY 521 TCATACAAACCTCCGCCATCTACT--GTATTTTTTTTAAATGAATATGGAATGATAAT 578
 Db 75535 TTGCTACAAACCTGACCCATTTGCTACATCTTTTTTCTATGAATATGGAATGCAAT 75594
 QY 579 A 579
 Db 75595 A 75595
 RESULT 9
 LOCUS AY077636 368 bp mRNA linear PRI 12-FEB-2002
 DEFINITION Homo sapiens zinc-induced metallothionein 2A/1A hybrid mRNA, complete cds.
 ACCESSION AY077636
 VERSION AY077636
 KEYWORDS AY077636.1 GI:18652805
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 368)
 AUTHORS Van Weyenbergh, J. and De Ley, M.
 TITLE Human monocyte zinc-induced mRNA MT2A/1A hybrid
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 368)
 AUTHORS Van Weyenbergh, J. and De Ley, M.
 TITLE Direct Submission
 JOURNAL Submitted (29-JAN-2002) LIM1-CPqGM, FIOCRUZ, Rua Waldemar Falcão 121, Salvador, BA 40295-001, Brazil
 FEATURES
 source
 1..368
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ZAP/B2"
 /cell_type="monocyte"
 /country="Belgium"
 14..199
 /codon_start=1
 /product="zinc-induced metallothionein 2A/1A hybrid"
 /protein_id="AA176084.1"
 /db_xref="GI:18652806"
 /translation="MDPNCSCAAGDSCTCAGSKCKECKTCKSKSCSCGCPMSCAKA
 AQGCIKGAASEKSCCA"
 BASE COUNT 106 a 98 c 78 g 86 t
 ORIGIN
 Query Match 44.6%; Score 259.8; DB 9; Length 368;
 Best Local Similarity 89.0%; Pred. No. 7.6e-62;
 Matches -298; Conservative 0; Mismatches 22; Indels 15; Gaps 1;
 QY 260 TCCTTGCTCGAAATGGACCCCACTGCTCTCGCTCGCTGTGGCTCTGCTGCTGCTGCC 319
 Db 1 TCCTGAGCTCGCAATGGATGCCAACTGCTCTCGCGCGCGGTGACTCTGCACCTGGCG 60
 QY 320 CGGCTCTGCAATGCAAGAGTGCNAATGCACTCTCTGCAAGAGAGTGCCTGCTCTG 379
 Db 61 CGGCTCTGCAATGCAAGAGTGCNAATGCACTCTCTGCAAGAGAGTGCCTGCTCTG 120
 QY 380 CTGCCCTGTGGCTGTGCCAATGTGCCAGGGCTGCATCTGCAAGGGGCATCAGAA 439
 Db 121 CTGCCCATGAGCTGTGCCAATGTGCCAGGGCTGCATCTGCAAGGGGCATCAGAA 180
 QY 440 GTGCAGCTCTGCTGTGATGTCGGACAGCCCTGCTCGAAGATATAGAAAGTGCCT 499
 Db 181 GTGCAGCTCTGCTGTGATGTCGGACAGCCCTGCTCGAAGATATAGAAAGTGCCT 240
 QY 500 GCACAACCTTGAATTTTTTTT-----CCATACACCCCTGCCCATCTAC 544
 Db 241 GCACAACCTTGAATTTTTTTTCCATACACCCCTGACCCATACACCCCTGACCCATTAC 300
 QY 545 TGTATTTTTTTTAAATGAATATGTGAATGATAATA 579

Db 301 TGTATTTTTTTTAAATGAATATGTGAATGATAATA 335
 RESULT 10
 LOCUS HS1174N9 72930 bp DNA linear PRI 17-MAY-2000
 DEFINITION Human DNA sequence from clone RP5-1174N9 on chromosome 1p34.1-35.3. Contains the gene for a novel protein with IBR domain, a (pseudo?) gene for a novel protein similar to MTE (metallothionein 1E (functional)), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.
 ACCESSION AL031502
 VERSION AL031502.14 GI:6729581
 KEYWORDS HTG; CpG Island; IBR; metallothionein; MTE.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 72930)
 AUTHORS Ellington, A.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT
 On Jan 21, 2000 this sequence version replaced gi:6634975.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, ENBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1
 RP5-1174N9 is from the library RPCI-5 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/VECTOR: pcypac2
 This sequence is the entire insert of clone RP5-1174N9.
 FEATURES
 Location/Qualifiers
 1..72930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="p34.1-35.3"
 /clone="RP5-1174N9"
 /clone_lib="RPCI-5"
 repeat_region 8..123
 /note="MIR repeat: matches 66. .181 of consensus"
 repeat_region 124..418
 /note="AluX repeat: matches 1. .297 of consensus"
 repeat_region 419..477
 /note="MIR repeat: matches 181. .239 of consensus"
 repeat_region 486..532
 /note="Alu repeat: matches 248. .294 of consensus"
 misc_feature 516..520
 /note="One clone deleted 5 base pairs."
 repeat_region 536..652
 /note="AluSg/x repeat: matches 1. .117 of consensus"
 repeat_region 663..827

Db 262 GAGTCAAAATGACCTCTGCAAGAAGAGCTGCTCCNTGCTCCNTGCTGGGCTGTC 203
QY 398 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 457
Db 202 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 143
QY 458 ATGTCCGGACAGCCCTGCTGCAAGATATAGAAAGAGTGCAGCAAACTTGA---A 513
Db 142 ATGCCAGCAGCTGCTGCTCAGATGTAATAGAGCAACCTATATAAACCTGGATT 83
QY 514 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571
Db 82 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 23
QY 572 TGAATA 579
Db 22 TGGCAATA 15

RESULT 13
LOCUS AX331811/c 409 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2320 from Patent WO0194629.
ACCESSION AX331811
VERSION AX331811.1 GI:18122445
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2320 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1. 409
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 116 a 93 c 109 g 87 t 4 others
ORIGIN

Query Match 42.1%; Score 245.2; DB 6; Length 409;
Best Local Similarity 88.0%; Pred. No. 9.6e-58;
Matches 324; Conservative 0; Mismatches 35; Indels 9; Gaps 5;

QY 221 CCGTGTGCTGTTTCCCTCTG-ATCGGGAATCCTGCTGCTGCTGCTGCTGCTGCTG 279
Db 382 CCGTGTGCTGTTTCCCTCTG-ATCGGGAATCCTGCTGCTGCTGCTGCTGCTGCTG 323
QY 280 CCAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 338
Db 322 CCAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263
QY 339 GAGTGCATAATGACCTCTGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397
Db 262 GAGTGCATAATGACCTCTGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203
QY 398 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 457
Db 202 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 143
QY 458 ATGTCCGGACAGCCCTGCTGCAAGATATAGAAAGAGTGCAGCAAACTTGA---A 513
Db 142 ATGCCAGCAGCTGCTGCTCAGATGTAATAGAGCAACCTATATAAACCTGGATT 83
QY 514 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571
Db 82 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 23
QY 572 TGAATA 579
Db 22 TGGCAATA 15

RESULT 15
LOCUS AX336280/c 409 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6789 from Patent WO0194629.
ACCESSION AX336280
VERSION AX336280.1 GI:18126999
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

Db 22 TGGCAATA 15

RESULT 14
LOCUS AX336059 409 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6568 from Patent WO0194629.
ACCESSION AX336059
VERSION AX336059.1 GI:18126778
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6568 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1. 409
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 116 a 93 c 109 g 87 t 4 others
ORIGIN

Query Match 42.1%; Score 245.2; DB 6; Length 409;
Best Local Similarity 88.0%; Pred. No. 9.6e-58;
Matches 324; Conservative 0; Mismatches 35; Indels 9; Gaps 5;

QY 221 CCGTGTGCTGTTTCCCTCTG-ATCGGGAATCCTGCTGCTGCTGCTGCTGCTGCTG 279
Db 382 CCGTGTGCTGTTTCCCTCTG-ATCGGGAATCCTGCTGCTGCTGCTGCTGCTGCTG 323
QY 280 CCAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 338
Db 322 CCAACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263
QY 339 GAGTGCATAATGACCTCTGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397
Db 262 GAGTGCATAATGACCTCTGCAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203
QY 398 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 457
Db 202 CAAGTGTCCAGGCTGCTATCGCAAGGGGATCAGAGAGTGCAGCTGCTGTCCTG 143
QY 458 ATGTCCGGACAGCCCTGCTGCAAGATATAGAAAGAGTGCAGCAAACTTGA---A 513
Db 142 ATGCCAGCAGCTGCTGCTCAGATGTAATAGAGCAACCTATATAAACCTGGATT 83
QY 514 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571
Db 82 TTTTTCATACAACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 23
QY 572 TGAATA 579
Db 22 TGGCAATA 15

RESULT 15
LOCUS AX336280/c 409 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6789 from Patent WO0194629.
ACCESSION AX336280
VERSION AX336280.1 GI:18126999
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6789 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
Source 1..409
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 116 a 93 c 109 g 87 t 4 others
ORIGIN
Query Match 42.1%; Score 245.2; DB 6; Length 409;
Best Local Similarity 88.0%; Pred. No. 9.6e-58;
Matches 324; Conservative 0; Mismatches 35; Indels 9; Gaps 5;
QY 221 CGCTCGGTGTTTCTCTTG-ATCGGGAACCTCTCTTCTCTTGCCTCGAAATGGACC 279
Db |||||
382 CGCTCGGTGTTTCTCTTGAATCGGAACCTCTCTTCTCTTGCCTCGAAATGGACC 323
QY 280 CCAACTGCTCCCTGC-TCGCTGTGTGCTCTCTGTGCTGTGCGGCTCCTGCAAAATGCAAA 338
Db |||||
322 CCAACTGCTCCCTGCNTCGCTGTGCTCTCTGTGCTGTGCGGCTCCTGCAAAATGCAAA 263
QY 339 GAGTGCAAATGCACCTCCTGCAAGAAGAGCTGCTGCC-TGCTGCCCTGTGGGCTGTGC 397
Db |||||
262 GAGTGCAAATGCACCTCCTGCAAGAAGAGCTGCTGCCNTGCTGTGGGCTGTGC 203
QY 398 CAAGTGTGCCAGGGCTGCATCTGCAAGGGGATCAGAGAGTGCAGCTGTGCTGTGCCCTG 457
Db |||||
202 CAAGTGTGCCAGGGCTGCATCTGCAAGGGGATCAGAGAGTGCAGCTGTGCTGTGCCCTG 143
QY 458 ATGTCGGGACAGCCCTGCTCGAAGATATAGAAGAGTGCCTGCACAAACTTGGG- ---A 513
Db |||||
142 ATGCCAGGACAGCTGTGCTCTCAGATGTAAATAGAGCAACCTATATAAACCTGGATTTT 83
QY 514 TTTTTCATACAAACCTGCCCCATCTACT--GTATTTTTCATGAAATATGTGAA 571
Db |||||
82 TTTTTCATACAAACCTGCCCCATCTACT--GTATTTTTCATGAAATATGTGAA 23
QY 572 TGATAATA 579
Db |||||
22 TGGCAATA 15

Search completed: July 16, 2003, 13:27:55
Job time : 1063.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 03:41:41 ; Search time 99.5867 Seconds
(without alignments)
13183.634 Million cell updates/sec

Title: US-09-981-353-114
Perfect score: 583
Sequence: 1 cccgtaggtgctgctctg.....tatgtgaatgataatagta 583

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	277.4	47.6	415	24	ABN97299 Gene #3797 used to
2	277.4	47.6	415	24	ABK64825 Human benign prost
3	277.4	47.6	415	24	ABL62482 Colon adenocarcino
4	277.4	47.6	415	24	ABL68231 Kidney cancer rela
5	270.6	46.4	448	22	ABL68578 Human polynucleoti
6	248	42.5	573	22	AA193455 Human polynucleoti
7	247	42.4	446	22	AA161205 Human polynucleoti
8	247	42.4	601	22	AA159419 Human polynucleoti
9	245.2	42.1	409	24	AAK52094 Gene #2703 used to
c					ABN96205

c	10	245.2	42.1	409	24	ABK64607	Human benign prost
c	11	245.2	42.1	409	24	ABL62407	Colon adenocarcino
c	12	245.2	42.1	409	24	ABL63983	Breast cancer rela
c	13	245.2	42.1	409	24	ABL68231	Kidney cancer rela
c	14	245.2	42.1	409	24	ABL68452	Kidney cancer rela
c	15	245.2	42.1	409	24	ABL68852	Kidney cancer rela
c	16	232.8	39.9	453	21	AAF16386	Human prostate can
c	17	227.6	39.0	682	21	AAF16055	Human prostate can
c	18	227.2	39.0	420	22	AA193800	Human polynucleoti
c	19	223.6	38.4	389	22	AA45033	CDNA encoding nove
c	20	221.8	38.0	641	22	AA523089	DNA encoding novel
c	21	219.8	37.7	363	16	AAAT20036	Human gene signatu
c	22	219.8	37.7	367	24	ABN97265	Gene #3763 used to
c	23	219.8	37.7	367	24	ABK64818	Human benign prost
c	24	219.8	37.7	367	24	ABL68636	Kidney cancer rela
c	25	217.4	37.3	353	16	AAAT20977	Human gene signatu
c	26	216.8	37.2	557	22	AAK51621	Human polynucleoti
c	27	216.4	37.1	500	21	AAAC99226	Human pancreatic c
c	28	216	37.0	646	22	AAF44864	Human breast cance
c	29	216	37.0	673	22	AAH30041	TRO005 kappa chain
c	30	216	37.0	678	22	AAH41664	Human interleukin
c	31	216	37.0	678	22	AAH41667	Human interleukin
c	32	216	37.0	678	22	AAH41670	Human interleukin
c	33	216	37.0	678	22	AAH30010	Anti-IL8 monoclonal
c	34	216	37.0	678	22	AAH30013	Anti-IL8 monoclonal
c	35	216	37.0	678	22	AAH30016	Anti-IL8 monoclonal
c	36	216	37.0	923	21	AA95785	Human immune syste
c	37	216	37.0	944	22	AA44892	Human breast cance
c	38	216	37.0	948	24	AA599475	Anti-human ALIM m
c	39	216	37.0	958	22	AAAC66520	Human immune syste
c	40	216	37.0	970	24	AA599477	Anti-human ALIM m
c	41	214.4	36.8	645	16	AAQ67896	CDNA contg. an ORF
c	42	214.4	36.8	678	22	AAH41652	Human interleukin
c	43	214.4	36.8	678	22	AAH29998	Anti-IL8 monoclonal
c	44	214.4	36.8	726	17	AAH18060	Monoclonal antibod
c	45	214.4	36.8	758	20	AAAX01213	Human antiFc eps11

ALIGNMENTS

RESULT 1
ABN97299
ID ABN97299 standard; DNA; 415 BP.
XX
AC ABN97299;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3797 used to diagnose liver cancer.
XX
KW Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
XX
PN WO200229103-A2.
XX
XX
PD 11-APR-2002.
XX
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
XX
PR 02-OCT-2000; 2000US-237054P.
XX
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,

expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Query Match	42.1%	Score 245.2	DB 24	Length 409
Best Local Similarity	88.0%	Pred. No. 1.5e-59		
Matches 324	Conservative 0	Mismatches 35	Indels 9	Gaps 5
QY	221	CCGCTGGGTCTTTTCCCTTGG - ATCGGGAACCTCTGCTTCTCTTCCTTGCCTCGAAATGGACC	279	
Db	382	CCGCTGGGTCTTTTCCCTTGAATCGGGAACCTCTGCTTCTCCTTGGCTCGAAATGGACC	323	
QY	280	CCAACRGCTCCTGC - TGGCCTGTTGGCTTCTGTGCGCTGTGCGCGCTCCTTGCAAATGCAAA	338	
Db	322	CCAACRGCTCCTGCNTGCGCTGTTGGCTTCTGTGCGCTGTGCGCGCTCCTTGCAAATGCAAA	263	
QY	339	GAGTGCAAATGCACTCCCTCGAAGACAGCTGCTGCTCC - TGTGTGCCCTGTGGGCTGTGC	397	
Db	262	GAGTGCAAATGCACTCCCTCGAAGACAGCTGCTGCTCCNTGCTGTTGGGCTGTGC	203	
QY	398	CAAGTGTGCCAGGGCTGCATCTGCAAGGGGCATCAGAGAAGTGCAGCTGCTGTGCCGTG	457	
Db	202	CAAGTGTGCCAGGGCTGCATCTGCAAGGGACGTCAGACAAGTGCAGCTGCTGTGCCGTG	143	
QY	458	ATGTCCGGACAGCCCTGCTCGAAGATATAGAAGAGTGCACCTGCACAAACTTGA - - - - A	513	
Db	142	ATGCCAGGACAGCTGTGCTCTCAGATGTAAATAGAGCAACCTATATAAACCTGGATTTTT	83	
QY	514	TTTTTTTTTCCATACAACCCGTGCCCATCTACT - - GTATTTTTTTTTTAATGAAATATGTAA	571	
Db	82	TTTTTTTTTTGTACAACCCGTGACCCGTTTGGCTACATCTTTTTTCTATGAAATATGTAA	23	
QY	572	TGATAATA	579	
Db	22	TGGCAATA	15	

RESULT 15	
ABL68852/c	
ID	ABL68852 standard; DNA; 409 BP.
XX	
XX	ABL68852;
XX	
XX	15-MAY-2002 (first entry)
DT	
XX	Kidney cancer related gene sequence SEQ ID NO:7189.
DE	
XX	
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW	gene; ds.
XX	
XX	Homo sapiens.
OS	
XX	
PN	WO200194629-A2.
XX	
XX	13-DEC-2001.
PD	
XX	
XX	30-MAY-2001; 2001WO-US10838.
PF	
XX	
XX	05-JUN-2000; 2000US-209473P.
PR	05-JUN-2000; 2000US-209531P.
PR	18-SEP-2000; 2000US-233133P.
PR	

PR	18-SEP-2000;	2000US-233617P.
PR	20-SEP-2000;	2000US-234009P.
PR	20-SEP-2000;	2000US-234034P.
PR	20-SEP-2000;	2000US-234052P.
PR	22-SEP-2000;	2000US-234509P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235963P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237395P.
PR	02-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.
XX	(AVAL-)	AVALON PHARM

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI: 2002-188264/24.

WPI: 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1: SEQ ID 7189; 44pp: English.

Claim 1: SEQ ID 7189; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 409 BP: 116 A; 93 C; 109 G; 87 T; 4 other;

Sequence 409 BP; 116 A; 93 C; 109 G; 87 T; 4 other;

Query Match 42.1%; Score 245.2; DB 24; Length 409;
Best Local Similarity 88.0%; Pred. No. 1.5e-59;
Matches 324; Conservative 0; Mismatches 35; Indels 9; Gaps 5;

Qy	221	CCGCTGGGTGTTTTCCTCTTG-ATCGGGAACCTCCTGCTTCTCCTTGCCCTCGAAATGGACC	279
Db	382	CCGCTGGGTGTTTTCCTCTTGAAATCGGGAACCTCCTGCTTCTCCTTGCCCTCGAAATGGACC	323
Qy	280	CCAACTGCTCCTGC-TGCGCTGTGGCTCCTGTGCTGTGCGGCTCCTGCAAAATGCAAA	338
Db	322	CCAACTGCTCCTGCNTGCGCTGTGGCTCCTGTGCGGCTCCTGCAAAATGCAAA	263
Qy	339	GAGTGCAAAATGCACCTCCTGCAAGAGAGCTGTGCTCC-TGCTGCCCTGTGGGCTGTGC	397
Db	262	GAGTGCAAAATGCACCTCCTGCAAGAGAGCTGTGCTCCNTGCTGTGGGCTGTGC	203
Qy	398	CAAGTGTGCCAGGGCTGCAATCTGCAAGGGGATCAGAGAGTGCAGCTGTGCTGTGCCTG	457
Db	202	CAAGTGTGCCAGGGCTGCAATCTGCAAGGGGATCAGAGAGTGCAGCTGTGCTGTGCCTG	143
Qy	458	ATGTCGGGACAGCCCTGCTGCAAGATATAGAAAGAGTGACCTGCACAACTTGA- - - A	513
Db	142	ATGCCAGGACAGCTGTGCTCTCAGATGTAATAGACACCTATATAAACCTGGATTTT	83
Qy	514	TTTTTTTCCATACAAACCTGCCCCATCTACT- - GTATTTTTTTTAAATGAAATATGTGAA	571
Db	82	TTTTTTTTTTGTACAAACCTGACCCGTTTGCCTACATCTTTTTTCTATGAATATGTGAA	23
Qy	572	TGATAATA 579	
Db	22	TGCCAATA 15	

Search completed: July 16, 2003, 11:09:59
Job time : 101.587 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 08:51:42 ; Search time 636.456 Seconds
(without alignments)
14835.236 Million cell updates/sec

Title: US-09-981-353-114
Perfect score: 583
Sequence: 1 cctgtaggctgtctcttg.....tatgtgaatgataatagta 583

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST: *
1:  em_estba: *
2:  em_esthum: *
3:  em_estin: *
4:  em_estnu: *
5:  em_estov: *
6:  em_estpl: *
7:  em_estro: *
8:  em_htc: *
9:  gb_est1: *
10: gb_est2: *
11:  gb_htc: *
12:  gb_est3: *
13:  gb_est4: *
14:  gb_est5: *
15:  em_estfun: *
16:  em_estom: *
17:  gb_gss: *
18:  em_gss_hum: *
19:  em_gss_inv: *
20:  em_gss_pln: *
21:  em_gss_vrt: *
22:  em_gss_fun: *
23:  em_gss_mam: *
24:  em_gss_mus: *
25:  em_gss_other: *
26:  em_gss_pro: *
27:  em_gss_rod: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	304.4	52.2	376	14	BM694421
2	292.4	50.2	516	12	BG566970
3	286.8	49.2	406	10	AW952177
4	282.8	48.5	399	14	BQ440401
5	280.8	48.2	499	13	BM458221
6	279.8	48.0	410	14	BQ428187

7	278.2	47.7	1509	14	BM924728
8	276.8	47.5	513	12	BG900215
9	276.8	47.5	568	13	BG926575
10	276.8	47.5	575	13	BG926326
11	276.8	47.5	1074	14	BM921692
12	276.8	47.5	1159	13	BM544354
13	275.8	47.3	493	12	BG897617
14	275.2	47.2	576	13	BG925124
15	274.8	47.1	407	10	AV661102
16	274.8	47.1	439	10	AV660814
17	274.8	47.1	639	13	BG925334
18	274.6	47.1	352	14	F32133
19	273.6	46.9	467	12	BF526059
20	273.2	46.9	618	13	BG928106
21	272.4	46.7	445	12	BG899081
22	272.2	46.7	518	13	BG927843
23	271.8	46.6	705	12	BF031192
24	271.4	46.6	828	12	BG536644
25	270.8	46.4	414	14	F26428
26	270.6	46.4	323	14	F26407
27	270.6	46.4	454	12	BG564723
28	270.6	46.4	563	12	BF676799
29	270.6	46.4	722	12	BG566693
30	270.6	46.4	744	12	BG613385
31	270.6	46.4	917	12	BF130769
32	270	46.3	734	13	BF159852
33	270	46.2	935	12	BG616113
34	269.6	46.2	426	12	BG900998
35	269	46.1	958	12	BG533343
36	268.2	46.0	503	10	AV662264
37	267.2	45.8	1186	12	BG568840
38	266.4	45.7	976	12	BG540271
39	265.8	45.6	332	14	BM847755
40	265.6	45.6	301	14	F27787
41	264.4	45.4	442	13	BG927047
42	263.2	45.1	317	14	F26137
43	262.6	45.0	315	14	F24837
44	261.2	44.8	296	14	F30476
45	261.2	44.8	461	12	BF766332

ALIGNMENTS

RESULT 1
BM694421
LOCUS
DEFINITION
BM694421
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

BM694421 376 bp mRNA linear EST 28-FEB-2002
UI-E-Cil-afg-d-21-0-UI.r1 UI-E-Cil Homo sapiens cDNA clone
UI-E-Cil-afg-d-21-0-UI 5', mRNA sequence.
BM694421
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 376)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

**JOURNAL
COMMENT**

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12809 row: f column: 11
High quality sequence stop: 486.

FEATURES

```

I. .1509
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5761162"
/clone_lib="NIH_MGC_116"
/lab_host="DH10b"
/notes="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPORT6; Site_: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
494 a 450 c 264 g 298 t 3 others

```

RESULT 8	ACCESSION
BG900215	VERSION
LOCUS	
DEFINITION	

BG9000215 513 bp mRNA linear EST 06-NOV-2001
 HOA36-1-B9 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA,
 mRNA sequence.
 BG9000215
 BG9000215.1 GI:14310464

KEYWORDS
SOURCE
ORGANIS

EST.
human.
Homo sapiens

REFERENCES
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 513)
Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J.,
Sathe, G., Mu, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and
Lark, M.W.

TITLE	JOURNAL	MEDLINE	COMMENT
-------	---------	---------	---------

Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries
Osteoarthr. Cartil. 9 (7), 641-653 (2001)

MEDLINE COMMENT	FEATURES	SOURCE

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

Location/Qualifiers
1. .513
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HOA (Human Oste
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/note="vector: pSPORT I; SI
directional"

```

DIRECTIONAL			
BASE COUNT	141 a	130 c	110 g
ORIGIN			
132 t			

Query Ma
Best Loc

Query Match	47.5%	Score	276.8	DB 12	Length	513		
Best Local Similarity	88.4%	Pred. No.	6.6e-65					
Matches 327	Conservative 0	Mismatches 32	Indels 11	Gaps 2				
QY	221	CGCGTGC	CGTGTTCCTCTT	GATCGGA	ACTCCTTCCTCTTCCTT	GCCTCGA	AAATGCACCC	280
Db	85	CGCGTGC	GTGTTTCCTCTT	GATCGGA	ACTCCTTCCTCTTCCTT	GCCTCGA	AAATGCACCC	144
QY	281	CAACTGCT	CCCTGCTCGCCT	GTGTGGCTCCT	TGCGG	GCTCCTGCA	AAATGCA	340
Db	145	CAACTGCT	CCCTGCTCGCCT	GTGTGGCTCCT	TGCGG	GCTCCTGCA	AAATGCA	204
QY	341	GTGCAAA	TGCACTTCCTG	CAAGAGAGCT	GCTGCTGCTG	CGCCCTGTG	GGGTGTGCC	400
Db	205	GTGCAAA	TGCACTTCCTG	CAAGAGAGCT	GCTGCTGCTG	CGCCCTGTG	GGGTGTGCC	264
QY	401	GTGTGCC	CAGGGCTGC	ATCTGCA	AAAGGGCATCAG	AGAAATG	CGAGCTGCTG	560
Db	265	GTGTGCC	CAGGGCTGC	ATCTGCA	AAAGGGCATCAG	AGAAATG	CGAGCTGCTG	324
QY	461	TCCGG	CAGCCCTGCT	CGAAGATAT	AGAAAGATG	ACCTGC	ACAAACTT	512
Db	325	CCAGG	ACAGCTGTGCT	CTCAGATGT	AAATAG	ACCAACCTAT	ATAAACCT	384
QY	513	-ATTTT	TTTTTTCAT	FACAA	CCCTGCC	CCATCTACT	--GTAT	569
Db	385	TTTTT	TTTTTTCAT	FACAA	CCCTGCC	CCGTGTG	TACATCT	444
QY	570	AATGA	TAAATA	579				
Db	445	AATGC	CAATA	454				

RESULT 9
BG926575
LOCUS
DEFINITION

QY 513 -ATTTTTCATACAAACCTGCCCATCTACT--GTATTTTTCATGAATATGTG 569
 Db 304 TTTTTCATACAAACCTGCCCATCTACTTTTTCATGAATATGTG 363
 QY 570 AATGATAATA 579
 Db 364 AATGGCAATA 373

RESULT 13
 BG897617 493 bp mRNA linear EST 06-NOV-2001
 LOCUS HOA15-1-B10 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA,
 mRNA sequence.
 ACCESSION BG897617
 VERSION BG897617.1 GI:14307858
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 493)
 AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
 Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
 Lark,M.W.
 TITLE Identification and initial characterization of 5000 expressed
 sequenced tags (ESTs) each from adult human normal and
 osteoarthritic cartilage cDNA libraries
 JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 COMMENT Contact: Sanjay Kumar
 UW2109
 GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay_kumar-l@sk.com
 Seq primer: T7.

FEATURES
 source
 1..493
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HOA (Human Osteoarthritic Cartilage)"
 /tissue_type="cartilage"
 /lab_host="E.coli DH10 B"
 /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
 Directional"
 BASE COUNT 127 a 128 c 110 g 128 t
 ORIGIN

Query Match 47.3%; Score 275.8; DB 12; Length 493;
 Best Local Similarity 87.5%; Pred. No. 1.2e-64;
 Matches 328; Conservative 0; Mismatches 37; Indels 10; Gaps 2;
 QY 215 CTGAGCGCGTGGGTTCCTTGTGATCGGGAACCTCCTGCTTCCTTCCTTCGCTCGAAT 274
 Db 75 CACGCGTCCGCGGTGTTTCTCTTGTATCGGGAACCTCCTGCTTCCTTCGCTCGAAT 134
 QY 275 GGACCCCAACTGCTCCTGCTCGCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 334
 Db 135 GGACCCCAACTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
 QY 335 CAAAGAGTGAATGCACTCCCTGCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 394
 Db 195 CAAAGAGTGAATGCACTCCCTGCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254
 QY 395 TGCCAAGTGTCCCGAGGGCTGCTATCTGCAAGGGGCATCAGAGAAGTGCAGCTGCTGCTG 454
 Db 255 TGCCAAGTGTCCCGAGGGCTGCTATCTGCAAGGGGCATCAGAGAAGTGCAGCTGCTGCTG 314
 QY 455 CTGATGTCGCGACAGCCCTCTCGAAGATATAGAAAGAGTGACCTGCACAACTTGA-- 512
 Db 315 CTGATGCGGACAGAGCTGCTCTTCAGATGTAATAGAGCAACCTATATAAACCTGGATT 374

QY 513 -----ATTTTTCATACAAACCTGCCCATCTACT--GTATTTTTCATGAAT 564
 Db 375 TTTTTCATACAAACCTGCCCATCTACTTTTTCATGAAT 434
 QY 565 AATGATAATA 579
 Db 435 AATGATAATA 449

RESULT 14
 BG925124 576 bp mRNA linear EST 06-NOV-2001
 LOCUS HNC37-1-B5.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
 DEFINITION sequence.
 ACCESSION BG925124
 VERSION BG925124.1 GI:14319647
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 576)
 AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
 Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
 Lark,M.W.
 TITLE Identification and initial characterization of 5000 expressed
 sequenced tags (ESTs) each from adult human normal and
 osteoarthritic cartilage cDNA libraries
 JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 COMMENT Contact: Sanjay Kumar
 UW2109
 GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay_kumar-l@sk.com
 Seq primer: T7.

FEATURES
 source
 1..576
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HNC (Human Normal Cartilage)"
 /tissue_type="cartilage"
 /lab_host="E.coli DH10 B"
 /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
 Directional"
 BASE COUNT 150 a 150 c 129 g 147 t
 ORIGIN

Query Match 47.2%; Score 275.2; DB 13; Length 576;
 Best Local Similarity 88.1%; Pred. No. 1.9e-64;
 Matches 326; Conservative 0; Mismatches 33; Indels 11; Gaps 2;
 QY 221 CCGCTGCGTGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 280
 Db 90 CCGCTGCGTGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 149
 QY 281 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 340
 Db 150 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 209
 QY 341 GTGCAAAATGACCTTCCTGCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
 Db 210 GTGCAAAATGACCTTCCTGCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
 QY 401 GTGTGCCAGGGCTGCTATCTGCAAGGGGCATCAGAGAAGTGCAGCTGCTGCTGCTGCTG 460
 Db 270 GTGTGCCAGGGCTGCTATCTGCAAGGGGCATCAGAGAAGTGCAGCTGCTGCTGCTGCTG 329
 QY 461 TCCGGACAGCCCTGCTGCAAGATATAGAAAGAGTGCACCTGCACAACTTGA----- 512
 Db 1 TCCGGACAGCCCTGCTGCAAGATATAGAAAGAGTGCACCTGCACAACTTGA----- 512

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 10:35:27 ; Search time 22.7443 Seconds
(without alignments)
7860.974 Million cell updates/sec

Title: US-09-981-353-114
Perfect score: 583

Sequence: 1 cccgttaggtgctgctctg.....tatgtgaatgataatagtcga 593

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	216	37.0	958	3	US-09-049-672A-15
C 2	214.4	36.8	3819	4	US-09-042-353-393
C 3	214.4	36.8	3819	4	US-08-758-417A-243
C 4	212.8	36.5	2178	1	US-08-463-587A-24
C 5	212.8	36.5	2178	2	US-08-463-667A-2
C 6	212.8	36.5	2178	3	US-08-923-854-24
C 7	212.8	36.5	2178	5	PCT-US91-09133-25
C 8	212.8	36.5	7305	1	US-08-286-740-4
C 9	212.8	36.5	7305	5	PCT-US95-09576-4
C 10	212	36.4	468	1	US-08-236-311-11
C 11	212	36.4	468	3	US-08-457-918-11
C 12	212	36.4	642	1	US-08-157-101A-8
C 13	212	36.4	1066	1	US-08-157-101A-4
C 14	211.2	36.2	705	4	US-09-171-945-51
C 15	211.2	36.2	705	4	US-09-171-945-96
C 16	211.2	36.2	705	4	US-09-171-945-98
C 17	211.2	36.2	708	1	US-08-488-376-18
C 18	211.2	36.2	708	2	US-08-634-223-18
C 19	211.2	36.2	708	2	US-08-634-224-18
C 20	211.2	36.2	708	2	US-08-634-400-18
C 21	211.2	36.2	708	2	US-08-635-878-18
C 22	211.2	36.2	708	2	US-08-770-057-18
C 23	211.2	36.2	708	4	US-09-335-697B-18
C 24	211.2	36.2	708	4	US-09-335-697B-18
C 25	211.2	36.2	780	3	US-09-027-449-54
C 26	211.2	36.2	780	3	US-09-027-449-58
C 27	211.2	36.2	780	3	US-09-027-449-65

C 28	211.2	36.2	780	3	US-08-804-444A-54	Sequence 54, Appl
C 29	211.2	36.2	780	3	US-08-804-444A-58	Sequence 58, Appl
C 30	211.2	36.2	780	3	US-09-026-985-54	Sequence 54, Appl
C 31	211.2	36.2	780	3	US-09-026-985-58	Sequence 58, Appl
C 32	211.2	36.2	780	3	US-09-026-985-65	Sequence 65, Appl
C 33	211.2	36.2	780	4	US-09-121-952A-54	Sequence 54, Appl
C 34	211.2	36.2	780	4	US-09-121-952A-58	Sequence 58, Appl
C 35	211.2	36.2	780	4	US-09-121-952A-65	Sequence 65, Appl
C 36	211.2	36.2	780	4	US-09-234-340A-54	Sequence 54, Appl
C 37	211.2	36.2	780	4	US-09-234-340A-58	Sequence 58, Appl
C 38	211.2	36.2	780	4	US-09-234-340A-65	Sequence 65, Appl
C 39	211.2	36.2	1576	5	PCT-US95-11405-34	Sequence 34, Appl
C 40	211.2	36.2	2143	3	US-09-097-309-5	Sequence 5, Appl
C 41	211.2	36.2	2143	4	US-09-097-171A-9	Sequence 9, Appl
C 42	211.2	36.2	2143	4	US-09-460-587-5	Sequence 5, Appl
C 43	211.2	36.2	6127	2	US-08-887-352B-1	Sequence 1, Appl
C 44	211.2	36.2	6127	4	US-09-109-207C-1	Sequence 1, Appl
C 45	211.2	36.2	6127	4	US-09-296-005-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-049-672A-15/C
; Sequence 15, Application US/09049672A
; Patent No. 6135941
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry, Janice
; APPLICANT: Au-Young, Janice
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,672A
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0497 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ADENINB01

RESULT 8
US-08-286-740-4/c
; Sequence 4, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
; APPLICANT: Crowley, Craig W.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,740
; FILING DATE: 05-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 798
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7305 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
PCT-US95-09576-4
;
Query Match 36.5%; Score 212.8; DB 1; Length 7305;
Best Local Similarity 96.9%; Pred. No. 6.1e-55;
Matches 217; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CTGTAGGTGCTGTCCTTGTCTGTGACACTCTCTCTGGGAGTACCCGATTGG 62
Db 3416 CTGTAGGTGCTGTCCTTGTCTGTGACACTCTCTCTGGGAGTACCCGATTGG 3357

QY 63 AGGGCGTTATCCACCTTCCACTTCTCTGGGAGTACCCGATTGG 122
Db 3356 AGGGCGTTATCCACCTTCCACTTCTCTGGGAGTACCCGATTGG 3297

QY 123 CACACAACAGAGCAGCTTCCAGATTTCAACTGCTCATCAGATGCGGGAAGATGAAGACA 182
Db 3296 CACACAACAGAGCAGCTTCCAGATTTCAACTGCTCATCAGATGCGGGAAGATGAAGACA 3237

QY 183 GATGGTGCAGCCACAGTTGCTTTGATCTCCAGCTCGAGCGCGTG 226
Db 3236 GATGGTGCAGCCACAGTTGCTTTGATCTCCAGCTCGAGCGCGTG 3193

RESULT 9
PCT-US95-09576-4/c
; Sequence 4, Application PC/TUS9509576
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09576
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286740
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
```

```
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 798PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7305 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
PCT-US95-09576-4
;
Query Match 36.5%; Score 212.8; DB 5; Length 7305;
Best Local Similarity 96.9%; Pred. No. 6.1e-55;
Matches 217; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CTGTAGGTGCTGTCCTTGTCTGTGACACTCTCTCTGGGAGTACCCGATTGG 62
Db 3416 CTGTAGGTGCTGTCCTTGTCTGTGACACTCTCTCTGGGAGTACCCGATTGG 3357

QY 63 AGGGCGTTATCCACCTTCCACTTCTCTGGGAGTACCCGATTGG 122
Db 3356 AGGGCGTTATCCACCTTCCACTTCTCTGGGAGTACCCGATTGG 3297

QY 123 CACACAACAGAGCAGCTTCCAGATTTCAACTGCTCATCAGATGCGGGAAGATGAAGACA 182
Db 3296 CACACAACAGAGCAGCTTCCAGATTTCAACTGCTCATCAGATGCGGGAAGATGAAGACA 3237

QY 183 GATGGTGCAGCCACAGTTGCTTTGATCTCCAGCTCGAGCGCGTG 226
Db 3236 GATGGTGCAGCCACAGTTGCTTTGATCTCCAGCTCGAGCGCGTG 3193

RESULT 10
US-08-236-311-11/c
; Sequence 11, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
```


NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 444PIC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-236-311-11

Query Match 36.4%; Score 212; DB 1; Length 468;
Best Local Similarity 97.7%; Pred. No. 2.5e-55;
Matches 215; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CTGTAGTGCTGTCTCTGTGACACTCTCTGTGACACTCTCTGTGAGTACACCGATTGG 62
DB 314 CTGTAGTGCTGTCTCTGTGACACTCTCTGTGACACTCTCTGTGAGTACACCGATTGG 255
QY 63 AGGGCGTTATCCACCTTCCACTGTACTTTGGCCTCTCTGGGATAGAAATTATTACGACAG 122
DB 254 AGGGCGTTATCCACCTTCCACTGTACTTTGGCCTCTCTGGGATAGAAATTATTACGACAG 195
QY 123 CACACACAGAGGCGAGTCCAGATTTCAACTGCTCATCAGATGCGCGGAGATGAAGACA 182
DB 194 CACACACAGAGGCGAGTCCAGATTTCAACTGCTCATCAGATGCGCGGAGATGAAGACA 135
QY 183 GATGTGCGAGCCACAGTTCGTTTGATCTCCAGTCGAGCC 222
DB 134 GATGTGCGAGCCACAGTTCGTTTGATCTCCAGTCGAGCC 95

RESULT 11

US-08-457-918-11/c
Sequence 11, Application US/08457918
Patent No. 6117655
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
TITLE OF INVENTION: Gregory, Timothy J.
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444PIC3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-457-918-11

Query Match 36.4%; Score 212; DB 3; Length 468;
Best Local Similarity 97.7%; Pred. No. 2.5e-55;
Matches 215; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CTGTAGTGCTGTCTCTGTGACACTCTCTGTGACACTCTCTGTGAGTACACCGATTGG 62
DB 314 CTGTAGTGCTGTCTCTGTGACACTCTCTGTGACACTCTCTGTGAGTACACCGATTGG 255
QY 63 AGGGCGTTATCCACCTTCCACTGTACTTTGGCCTCTCTGGGATAGAAATTATTACGACAG 122
DB 254 AGGGCGTTATCCACCTTCCACTGTACTTTGGCCTCTCTGGGATAGAAATTATTACGACAG 195
QY 123 CACACACAGAGGCGAGTTCAGATTTCAACTGCTCATCAGATGCGCGGAGATGAAGACA 182
DB 194 CACACACAGAGGCGAGTTCAGATTTCAACTGCTCATCAGATGCGCGGAGATGAAGACA 135
QY 183 GATGTGCGAGCCACAGTTCGTTTGATCTCCAGTCGAGCC 222
DB 134 GATGTGCGAGCCACAGTTCGTTTGATCTCCAGTCGAGCC 95

RESULT 12

US-08-157-101A-8/c
Sequence 8, Application US/08157101A
Patent No. 5808032
GENERAL INFORMATION:
APPLICANT: KURIHARA, TATSUYA
APPLICANT: MATSUKURA, SHIGEKAZU
APPLICANT: TSURUOKA, NOBUO
APPLICANT: ARIMA, KENJI
APPLICANT: NISHIHARA, TATSURO
TITLE OF INVENTION: ANTI-HBs ANTIBODY GENES AND EXPRESSION
TITLE OF INVENTION: PLASMIDS THEREFOR
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/157,101A
FILING DATE: 05-APR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: TITUS, MARLANA K
REGISTRATION NUMBER: 35843
REFERENCE/DOCKET NUMBER: 9437/204199

RESULT 15

Query Match	36.2%	Score 211.2;	DB 4;	Length 705;
Best Local Similarity	96.4%	Pred. No. 5.4e-55;		
Matches 216;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
QY	3	CTGTAGTGCTGTCCCTTGCTGCTCTGTGTGACACTCTCTCGGGAGTCAACCGATTGG	62	
DB	584	CTGTAGTGCTGTCCCTTGCTGCTCTGTGTGACACTCTCTCGGGAGTTCACCGATTGG	525	
QY	63	AGGCGCTTATCCACCTTCCACTCTACTTTGGCCCTCTCTGGGATAGAAGTTATTACGACGG	122	
DB	524	AGGCGCTTATCCACCTTCCACTCTACTTTGGCCCTCTCTGGGATAGAAGTTATTACGACGG	465	
QY	123	CACACAACAGAGGCAGTTCACAGATTTCACACTGCTCATCAGATGCGGGAGATGAAGACA	182	
DB	464	CACACAACAGAGGCAGTTCACAGATTTCACACTGCTCATCAGATGCGGGAGATGAAGACA	405	
QY	183	GATGTTGCAGCCACAGTTCGTTTGATCTCTCAGCTCGAGCCGCTG	226	
DB	404	GATGTTGCAGCCACAGTTCGTTTGATCTCTCAGCTTCGTTGCCCTTG	361	

Search completed: July 16, 2003, 14:58:51
Job time : 23.7443 secs

241 GATCGGGAACCTCCTGCTTCTCCTCGAAATGAGCCCAACTGCTCCTGCTGCGCTG 300
301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
361 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
361 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
421 GCAAGGGGATCAGAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
421 GCAAGGGGATCAGAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
481 GATATAGAAAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
481 GATATAGAAAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
541 CTACTGTATTTTATTAATGAATATGTAATGATAATAGTCA 583
541 CTACTGTATTTTATTAATGAATATGTAATGATAATAGTCA 583

RESULT 2

US-10-084-817-179

; Sequence 179, Application US/10084817

; Publication No. US20030119009A1

; GENERAL INFORMATION:

; APPLICANT: Susan Stuart

; APPLICANT: Jed G. Nuchtern

; APPLICANT: Sharon E. Plon

; APPLICANT: Jason M. Shohet

; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION

; FILE REFERENCE: PA-0046 US

; CURRENT APPLICATION NUMBER: US/10/084.817

; CURRENT FILING DATE: 2002-02-25

; PRIOR APPLICATION NUMBER: 60/270,784

; PRIOR FILING DATE: 2001-02-23

; NUMBER OF SEQ ID NOS: 365

; SOFTWARE: PERL Program

; SEQ ID NO 179

; LENGTH: 583

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20030119009A1 2685676CB1

US-10-084-817-179

Query Match 100.0%; Score 583; DB 9; Length 583;
Best Local Similarity 100.0%; Pred. No. 7.5e-179;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 CCCTGTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

Qy 61 GGAGGGCGTTATCCACCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 GGAGGGCGTTATCCACCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Qy 121 GGCACACACAGAGGAGTCCAGATTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GGCACACACAGAGGAGTCCAGATTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 181 CAGATGTGAGCCACAGTTCGTTTATGATCTCCAGCTCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CAGATGTGAGCCACAGTTCGTTTATGATCTCCAGCTCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

Qy 241 GATCGGGAACCTCCTGCTTCTCCTCGAAATGAGCCCAACTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

241 GATCGGGAACCTCCTGCTTCTCCTCGAAATGAGCCCAACTGCTCCTGCTGCTGCTG 300
301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
361 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
361 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
421 GCAAGGGGATCAGAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
421 GCAAGGGGATCAGAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
481 GATATAGAAAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
481 GATATAGAAAGAGTGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
541 CTACTGTATTTTATTAATGAATATGTAATGATAATAGTCA 583
541 CTACTGTATTTTATTAATGAATATGTAATGATAATAGTCA 583

RESULT 3

US-10-044-090-584

; Sequence 584, Application US/10044090

; Patent No. US20020137081A1

; GENERAL INFORMATION:

; APPLICANT: Olga Bandman

; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

; FILE REFERENCE: PA-0028 US

; CURRENT APPLICATION NUMBER: US/10/044.090

; CURRENT FILING DATE: 2002-01-09

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: PERL Program

; SEQ ID NO 584

; LENGTH: 583

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20020137081A1 2685676CB1

US-10-044-090-584

Query Match 100.0%; Score 583; DB 12; Length 583;
Best Local Similarity 100.0%; Pred. No. 7.5e-179;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTGTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 CCCTGTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

Qy 61 GGAGGGCGTTATCCACCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 GGAGGGCGTTATCCACCTTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Qy 121 GGCACACACAGAGGAGTCCAGATTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 GGCACACACAGAGGAGTCCAGATTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 181 CAGATGTGAGCCACAGTTCGTTTATGATCTCCAGCTCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CAGATGTGAGCCACAGTTCGTTTATGATCTCCAGCTCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240

Qy 241 GATCGGGAACCTCCTGCTTCTCCTCGAAATGAGCCCAACTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GATCGGGAACCTCCTGCTTCTCCTCGAAATGAGCCCAACTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

Qy 301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 361 AGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420


```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incy
US-10-044-090-835

```

Query Match	Score	DB	Length
Best Local Similarity	40.28;	234.4;	12;
	82.18;	234.4;	12;

[illegible]

RESULT 14

```

US-09-925-300-821
; Sequence 821, Application US/09925300
; Patent NO. US20020151681A
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Protei
; FILE REFERENCE: P101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 821
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (392)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (409)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (430)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (433)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature

```

```

; LOCATION: (434)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-821

```

[illegible]

RESULT 15

```

US-09-918-995-8793
; Sequence 8793, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8793
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8793

```

	Query Match	39.1%	Score 228.2	DB 9	Length 409
	Best Local Similarity	83.5%	Pred. No. 9.8e-64		
	Matches 283	Conservative	0	Mismatches 53	Indels 3
				Gaps 2	
Qy	243	TCGGAACTCTGTCTTCCTGCCTCGAATGGACCCCACTGTCTCTGCGCTCTT	302		
Db	9	TGGGGAGCTCCAGATACCCCTTTCTCGAATGGACCCCACTGTCTCTGCGCCACTGGT	68		
Qy	303	GCCTCTCTGTGCGCTGTGCGGCTCCTGCAATGCAAGAGTGCAAAATGCACCTCTCTGCAAG	362		
Db	69	GGCTCTCTGACGTGCGCGGCTGCTGCAAGTGCAGAGTGCAAAATGCACCTCTCTGCAAG	128		
Qy	363	AAGAGCTCTCTCTCTCTGCTGCCCTGTGGCTGTGCCAAGTGTGCCAGGGTGCATCTGC	422		
Db	129	AAGAGCTGCGTTCCTCTCTGCCCGGTGGGCTGTGCCAAGTGTGCCAGGGTGCCTGTC	188		

Qy	423	AAAGGGCATCAGAGAGTGCAGCTGCTGCCCTGATGTCGGACAGCCCTGCTCGAAGA	482
Db	189	AAAGGGCATCGGAGAGTGCAGCTGCTGCCCTGATGTCGGACAGCTCTTCTCCCATATA	248
Qy	483	TATAGAAAGAGTGCAGCTGCACAAACTTGGAAATTTTTTCCATACAAACCCCTGCCCCCATCT	542
Db	249	TGTAATAGAACACCTGCAC-AACCTGGATTTTTTAAAAATACAACTGAGCCATTT	307
Qy	543	ACTGTATTTTTTTTAAAT--GAAATATGTGAATGATAATA	579
Db	308	GCTGCATTTCTTTTATATAATAATATGCTGACTGACAATA	346

Search completed: July 16, 2003, 18:04:19
Job time : 78.7139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: July 16, 2003, 03:49:46 ; Search time 2975.96 Seconds
(without alignments)
15998.962 Million cell updates/sec

Title: US-09-981-353-165
Perfect score: 1636
Sequence: 1 gatcagtggtgaggaact.....ccaaattcaagaagacotg 1636

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_hg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_on:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1636	100.0	1636	6	AX327326	Sequence
2	1636	100.0	2759	6	AX395164	Sequence
3	1636	100.0	2974	6	AX376454	Sequence
4	1444.4	88.3	2786	6	AK025587	Sequence
5	1020.4	62.4	2547	10	AF175221	Cavia por
6	967.4	59.1	1650	6	AX040087	Sequence
7	886	54.2	2103	10	BC025795	Mus muscu
8	821.2	50.2	1766	9	HSAG054	Sequence
9	782	47.8	2090	6	AX10846	Sequence
10	782	47.8	2090	6	HSU08854	Sequence
11	778.8	47.6	2114	9	AF180322	Sequence
12	775.6	47.4	2107	6	AF180322	Sequence
13	775.6	47.4	2107	6	AF180322	Sequence
14	767.6	46.9	2108	9	HSU59209	Sequence
15	766.2	46.8	2123	6	AX10846	Sequence
16	766.2	46.8	2123	6	AX10846	Sequence
17	766	46.8	2123	9	HSU06641	Sequence
18	764.4	46.7	2079	9	HSAGJ5162	Sequence
19	762.8	46.6	1584	10	AF184901	Sequence
20	762.8	46.6	1584	10	AF184901	Sequence
21	758	46.3	2091	9	AF081793	Sequence
22	758	46.3	2093	6	AX111198	Sequence
23	758	46.3	2093	6	AX111198	Sequence
24	757.6	46.3	2075	9	AF072223	Sequence
25	754.8	46.1	1722	9	AF016492	Sequence
26	754.8	46.1	1855	6	AX336329	Sequence
27	754.8	46.1	1855	6	AX336329	Sequence
28	754.8	46.1	1855	6	AX336329	Sequence
29	754.8	46.1	1855	6	AX409473	Sequence
30	751	45.9	1753	9	HUMDPGTA	Sequence
31	750	45.8	1753	9	AF016310	Sequence
32	748.4	45.7	1766	9	BC030974	Sequence
33	745.4	45.6	1915	9	AFU91582	Sequence
34	745.2	45.6	1662	9	AF401657	Sequence
35	743.4	45.4	217991	9	AC021146	Sequence
36	743	45.4	1831	10	RRUGTGENE	Sequence
37	738.8	45.2	1662	9	AF294902	Sequence
38	736	45.0	1851	9	AF177272	Sequence
39	734	44.9	1768	9	AF112113	Sequence
40	732.6	44.8	2799	6	AX111112	Sequence
41	732.6	44.8	2799	9	HSUGT2BIO	Sequence
42	732.2	44.8	21000	6	AX395166	Sequence
43	731.8	44.7	1671	4	RABUGT2BB	Sequence
44	717.2	43.8	1825	10	AB034987	Sequence
45	701.4	42.9	1832	4	RABUGT2BA	Sequence

ALIGNMENTS

RESULT 1
AX327326
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX327326
Sequence
AX327326
AX327326.1
GI:18097872
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Policky,J.L., Hafalia,A., Burford,N., Ring,H.2., Lal,P.,
Tribouley,C.M., Yao,M.G., Yue,H., Tang,Y.T., Patterson,C., Das,D.,
Sanjanwala,M.S., Gandhi,A.R., Reddy,R., Khan,F.A., Baughn,M.R.,

AX327326
Sequence
AX327326
AX327326.1
GI:18097872
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Policky,J.L., Hafalia,A., Burford,N., Ring,H.2., Lal,P.,
Tribouley,C.M., Yao,M.G., Yue,H., Tang,Y.T., Patterson,C., Das,D.,
Sanjanwala,M.S., Gandhi,A.R., Reddy,R., Khan,F.A., Baughn,M.R.,

1636 bp
DNA
linear
PAT 07-JAN-2002

Rankumar,J., Griffin,J.A. and Au-Young,J.

Drug metabolizing enzymes

Patent: WO 0179468-A 11 25-OCT-2001;

Incyte Genomics, Inc. (US)

Location/Qualifiers

source

1. 1636

/organism="Homo sapiens"

/db_xref="taxon:9606"

/note="Incyte ID No: 2434655CBI"

BASE COUNT 483 a 315 c 369 g 469 t

ORIGIN

Query Match 100.0%; Score 1636; DB 6; Length 1636;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTGC 60
Db 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTGC 60
Qy 61 TCCTGCAGCTCTCTGTGTGGCTGTGGATTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG 120
Db 61 TCCTGCAGCTCTCTGTGTGGCTGTGGATTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG 120
Qy 121 ACATGAGCCATTTGGCTTAATGTCAAGGTCAATCTTAGAAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTTGGCTTAATGTCAAGGTCAATCTTAGAAGAGCTCATAGTGAGAGGCCATG 180
Qy 181 AGGTAAACAGTATTGACTCACTCAAGCCCTTCGTTAATGTGACTACAGAGCCCTCTGCAT 240
Db 181 AGGTAAACAGTATTGACTCACTCAAGCCCTTCGTTAATGTGACTACAGAGCCCTCTGCAT 240
Qy 241 TGAATTTGAGTGGTCCATATGCCAGGACAGACAGAAATGAAATATTGTGTG 300
Db 241 TGAATTTGAGTGGTCCATATGCCAGGACAGACAGAAATGAAATATTGTGTG 300
Qy 301 ACCTAGCTGGAATGTCTGCGAGGCTTATCAAGCTGGCAATCAGTTATATAATTAATG 360
Db 301 ACCTAGCTGGAATGTCTGCGAGGCTTATCAAGCTGGCAATCAGTTATATAATTAATG 360
Qy 361 ATTTTGTGTAATTAAGAGAACTTTAAATGATGTGTGAGAGCTTTATCTACAAATC 420
Db 361 ATTTTGTGTAATTAAGAGAACTTTAAATGATGTGTGAGAGCTTTATCTACAAATC 420
Qy 421 AGACGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTATAGACCCCTGTA 480
Db 421 AGACGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTATAGACCCCTGTA 480
Qy 481 TTCCCTGTGAGAGCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGAGAGCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
Qy 541 TTTCTGTAGAGGCAATATGAGGCAAGCTGTGGAAACTTCCAGCTCCACTTTTCCCTATG 600
Db 541 TTTCTGTAGAGGCAATATGAGGCAAGCTGTGGAAACTTCCAGCTCCACTTTTCCCTATG 600
Qy 601 TACCTGTGCTTATGACAGGACTTAACAGACAGAATGACCTTTCTGGAAGAGTAAATAAT 660
Db 601 TACCTGTGCTTATGACAGGACTTAACAGACAGAATGACCTTTCTGGAAGAGTAAATAAT 660
Qy 661 CAATGCTTTCAGTTTGTTCACCTTCCTGAGTTCAGGATTCAGGATTCATCTTTTGGGAAG 720
Db 661 CAATGCTTTCAGTTTGTTCACCTTCCTGAGTTCAGGATTCAGGATTCATCTTTTGGGAAG 720
Qy 721 AGTTTATAGTAAGGCAATAGGAGGCCACTACATATGTGAGACTGTGGGAAGCTG 780
Db 721 AGTTTATAGTAAGGCAATAGGAGGCCACTACATATGTGAGACTGTGGGAAGCTG 780
Qy 781 AGATATGGCTAATACGAACATATTGGGATTTTGAATTTTCTTCAACCATACCAACTTAAT 840
Db 781 AGATATGGCTAATACGAACATATTGGGATTTTGAATTTTCTTCAACCATACCAACTTAAT 840
Qy 841 TTGAGTTTGTGGAGGATTGACATGTAAACCTGCCAAAGCTTTGCCCTAAGGAATGGAAA 900

Db 841 TTGAGTTTGTGGAGATTGCACGTGTAACCCGCAAGCTTTGCCTAAGGAATGGAAA 900
Qy 901 ATTTTGTCCAGAGTTTCAGGGAAAGCTGATTTGTTGTTCTCTCTGGGTCACCTGTTTC 960
Db 901 ATTTTGTCCAGAGTTTCAGGGAAAGCTGATTTGTTGTTCTCTCTGGGTCACCTGTTTC 960
Qy 961 AAAATGTTACAGAAAGAAAGCTTAATATCATTTGCTTCAGCCCTTTGCCAGATCCACAGA 1020
Db 961 AAAATGTTACAGAAAGAAAGCTTAATATCATTTGCTTCAGCCCTTTGCCAGATCCACAGA 1020
Qy 1021 AGTGTGTTATGGAGGTACAAAGGAAAAAACCATCCACATTTAGGAGCCAACTACTCGGTGT 1080
Db 1021 AGTGTGTTATGGAGGTACAAAGGAAAAAACCATCCACATTTAGGAGCCAACTACTCGGTGT 1080
Qy 1081 ATGATTTGATACCCAGAAATCATCTTCTGTGTCATCCCAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTTGATACCCAGAAATCATCTTCTGTGTCATCCCAACCAAGCTTTTATCACTC 1140
Qy 1141 ATGCTGGAATGAATGGGATCTATGAAGCTATTTACCATGGGTCCTTATGTTGGAGTTC 1200
Db 1141 ATGCTGGAATGAATGGGATCTATGAAGCTATTTACCATGGGTCCTTATGTTGGAGTTC 1200
Qy 1201 CCATATTTGTTGATGACGCTTGATTAACATAGCTCAGTGAAGGCCAAAGGAGCAGCTGTAG 1260
Db 1201 CCATATTTGTTGATGACGCTTGATTAACATAGCTCAGTGAAGGCCAAAGGAGCAGCTGTAG 1260
Qy 1261 AAATAAACTTCAAAACTATGACAAGCGAAGATTTTACTGAGGCTTTGAGAACAGCTCATT 1320
Db 1261 AAATAAACTTCAAAACTATGACAAGCGAAGATTTTACTGAGGCTTTGAGAACAGCTCATT 1320
Qy 1321 CCGATTCTCTTATAAAGAGAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Db 1321 CCGATTCTCTTATAAAGAGAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Qy 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTGATCGGCGCCACAAAGAGGCCA 1440
Db 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTGATCGGCGCCACAAAGAGGCCA 1440
Qy 1441 AGCACCTGCGATCAGCTGCCCATGACCTGCTGCTGATCGAGTTTGTGATCGGCGCCACAAAGAGGCCA 1500
Db 1441 AGCACCTGCGATCAGCTGCCCATGACCTGCTGCTGATCGAGTTTGTGATCGGCGCCACAAAGAGGCCA 1500
Qy 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTGTTCACAAAATGTTTTTAT 1560
Db 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTGTTCACAAAATGTTTTTAT 1560
Qy 1561 TTTCTCTCAAAAATTTAATAAACTAGAAAGATAGAAAGAGGGAATAGATCTTTCCAA 1620
Db 1561 TTTCTCTCAAAAATTTAATAAACTAGAAAGATAGAAAGAGGGAATAGATCTTTCCAA 1620
Qy 1621 ATTCAAGAAAGACCTG 1636
Db 1621 ATTCAAGAAAGACCTG 1636

RESULT 2

AX395164

LOCUS

DEFINITION

AX395164

ACCESSION

VERSION

AX395164.1

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1

AUTHORS

TITLE

JOURNAL

AX395164 2759 bp DNA linear PAT 18-MAY-2002
Sequence 1 from Patent WO0218554.
AX395164
AX395164.1 GI:21066217
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Wei, H., Ye, J., Ketchum, K.A., di Francesco, V. and Beasley, E.M.
Isolated human drug-metabolizing proteins, nucleic acid molecules
encoding human drug-metabolizing proteins, and uses thereof
Patent: WO 0218554-A 1 07-MAR-2002;
Applera Corporation (US)

BASE COUNT	946 a	/db_xref="taxon:9606"	556 g	975 t	
ORIGIN					
Query Match	100.0%	Score 1636;	DB 6;	Length 2974;	
Best Local Similarity	100.0%	Pred. No. 0;			
Matches 1636;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGCAAGTCTGACGCTTTGGTATTCTGCG	60		
DB	1	GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGCAAGTCTGACGCTTTGGTATTCTGCG	60		
QY	61	TCCTGCAGCTCTTCTGTGTGTGGTGTGGATTTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG	120		
DB	61	TCCTGCAGCTCTTCTGTGTGTGGTGTGGATTTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG	120		
QY	121	ACATGAGCCCATTTGCTTAATGTCNAGGTCAATCTTAGAAGAGCTCATAGTGAAGGCCATG	180		
DB	121	ACATGAGCCCATTTGCTTAATGTCNAGGTCAATCTTAGAAGAGCTCATAGTGAAGGCCATG	180		
QY	181	AGGTAACAGTATTGACCTCACTCAAGGCTTCGTTAAATTGACTACAGGAAGCCCTTCTGCAT	240		
DB	181	AGGTAACAGTATTGACCTCACTCAAGGCTTCGTTAAATTGACTACAGGAAGCCCTTCTGCAT	240		
QY	241	TGAAATTTGAGGTGGTCCATATGCCACAGACAGAACAGAAATGAAATATTGTTG	300		
DB	241	TGAAATTTGAGGTGGTCCATATGCCACAGACAGAACAGAAATGAAATATTGTTG	300		
QY	301	ACCTAGCTTGAATGCTTCCAGGCTTATCAACCTGGCAATCAGTTTATAAATTAATG	360		
DB	301	ACCTAGCTTGAATGCTTCCAGGCTTATCAACCTGGCAATCAGTTTATAAATTAATG	360		
QY	361	ATTTTTTTGTTGAATAAGAGGAACCTTAAATAATGATGTGTGAGAGCTTTATCTACAAATC	420		
DB	361	ATTTTTTTGTTGAATAAGAGGAACCTTAAATAATGATGTGTGAGAGCTTTATCTACAAATC	420		
QY	421	AGAGCTTTATGAGAAGCTACAGAAACCACTACGATGTAAGTCTTATAGACCCCTGTA	480		
DB	421	AGAGCTTTATGAGAAGCTACAGAAACCACTACGATGTAAGTCTTATAGACCCCTGTA	480		
QY	481	TTCCCTCTGAGAGCTCATGGCTGAGTGTCTGCAGTCCCTTTTGTGCTCACACTTAGAA	540		
DB	481	TTCCCTCTGAGAGCTCATGGCTGAGTGTCTGCAGTCCCTTTTGTGCTCACACTTAGAA	540		
QY	541	TTTCTGTAGGAGCAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTCCCTATG	600		
DB	541	TTTCTGTAGGAGCAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTCCCTATG	600		
QY	601	TACCTGCTCATGACAGGACTAACACAGAGAATGACCTTTCTGGAAAGAGTAAATAATT	660		
DB	601	TACCTGCTCATGACAGGACTAACACAGAGAATGACCTTTCTGGAAAGAGTAAATAATT	660		
QY	661	CAATGCTTTTCAGTTTCTCCACTTTCGATTCAGGATTACGACTATCATTTTGGGAAG	720		
DB	661	CAATGCTTTTCAGTTTCTCCACTTTCGATTCAGGATTACGACTATCATTTTGGGAAG	720		
QY	721	AGTTTTATAGTAAGGCAATAGGAGGCCCACTACATTATGTGAGACTGTGGGAAAGCTG	780		
DB	721	AGTTTTATAGTAAGGCAATAGGAGGCCCACTACATTATGTGAGACTGTGGGAAAGCTG	780		
QY	781	AGATATGGCTTAATACGAACATATTGGGATTTTGAATTTCTCTCAACCATACCACTAACT	840		
DB	781	AGATATGGCTTAATACGAACATATTGGGATTTTGAATTTCTCTCAACCATACCACTAACT	840		
QY	841	TTGAGTTTGTGGAGGATTCACCTGTAAACCTGCAAGCTTTTGCCTTAAGGAATGAAA	900		
DB	841	TTGAGTTTGTGGAGGATTCACCTGTAAACCTGCAAGCTTTTGCCTTAAGGAATGAAA	900		
QY	901	ATTTTGTCCAGAGTTTCCAGGGAAGATGGTATTGTGGTGTCTCTGGGCTCACTGTTTC	960		
DB	901	ATTTTGTCCAGAGTTTCCAGGGAAGATGGTATTGTGGTGTCTCTGGGCTCACTGTTTC	960		
QY	961	AAATGTTACAGAGAAAGGCTTAATATCATTTGCTTCAGCCCTTGGCCAGATCCACAGA	1020		
DB	961	AAATGTTACAGAGAAAGGCTTAATATCATTTGCTTCAGCCCTTGGCCAGATCCACAGA	1020		
QY	1021	AGGTGTTTATGGAGGTACAAAGGAAAAAACCATCCACATTAGGAGCAATACTCGGCTGT	1080		
DB	1021	AGGTGTTTATGGAGGTACAAAGGAAAAAACCATCCACATTAGGAGCAATACTCGGCTGT	1080		
QY	1081	ATGATTTGGATACCCAGCAATGATCTTCTGTGTCATCCCAAAACCAAGCTTTTATCACTC	1140		
DB	1081	ATGATTTGGATACCCAGCAATGATCTTCTGTGTCATCCCAAAACCAAGCTTTTATCACTC	1140		
QY	1141	ATGTTGGAATGAATGGGATCTATGAAGCTATTATTACCATGGGCTCCCTATGTTGGAGTTC	1200		
DB	1141	ATGTTGGAATGAATGGGATCTATGAAGCTATTATTACCATGGGCTCCCTATGTTGGAGTTC	1200		
QY	1201	CCATATTTGTTGATCAGCTTTCATCAACATAGCTCATGAAAGCCCAAGGAGCAGCTCTAG	1260		
DB	1201	CCATATTTGTTGATCAGCTTTCATCAACATAGCTCATGAAAGCCCAAGGAGCAGCTCTAG	1260		
QY	1261	AAATAAACTTCAAAACTATGACAGCGAAGATTTTACTGAGGGCTTTTGAGAACAGTCAATTA	1320		
DB	1261	AAATAAACTTCAAAACTATGACAGCGAAGATTTTACTGAGGGCTTTTGAGAACAGTCAATTA	1320		
QY	1321	CCGATTTCTCTTATAAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG	1380		
DB	1321	CCGATTTCTCTTATAAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG	1380		
QY	1381	TAAAGCCCTTAGATCGAGCAGTCTTCTGATCGAGTTTGTCTCATCGCGCCACAAAGAGGCCA	1440		
DB	1381	TAAAGCCCTTAGATCGAGCAGTCTTCTGATCGAGTTTGTCTCATCGCGCCACAAAGAGGCCA	1440		
QY	1441	AGCACCTGCGATCAGCTGCCCATGACCTACCTGTTTCCAGCACTACTCTATAGATGTGA	1500		
DB	1441	AGCACCTGCGATCAGCTGCCCATGACCTACCTGTTTCCAGCACTACTCTATAGATGTGA	1500		
QY	1501	TTGGTTTCTGCTGACCTGTGTGCAACTGCTATATTCTTGTTCACAAAATGTTTTTAT	1560		
DB	1501	TTGGTTTCTGCTGACCTGTGTGCAACTGCTATATTCTTGTTCACAAAATGTTTTTAT	1560		
QY	1561	TTTCTCTCTCAAAATTTTAAATAAACTAGAAAGATAGAAAAGAGGGAATAGATCTTTCCAA	1620		
DB	1561	TTTCTCTCTCAAAATTTTAAATAAACTAGAAAGATAGAAAAGAGGGAATAGATCTTTCCAA	1620		
QY	1621	ATTCAAGAAAGACCTG 1636			
DB	1621	ATTCAAGAAAGACCTG 1636			
RESULT 4					
AK025587					
LOCUS	2786 bp mRNA linear PRI 29-SEP-2000				
DEFINITION	Homo sapiens cDNA: FLJ21934 fis, clone HEP04364.				
ACCESSION	AK025587				
VERSION	AK025587.1 GI:10438147				
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP clone:HEP04364.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (sites)				
TITLE	Kawabata,A., Hiki,j.,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.				
JOURNAL	NEDO human cDNA sequencing project				
REFERENCE	2 (bases 1 to 2786)				
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,				

Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

Location/Qualifiers

1..2786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP04364"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
71..1420
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15179.1"
/db_xref="GI:10438148"

translation="NPQDTEENEFVDLALNPLSLTWSQVILNDFVFEIRGLTK
MCESEIYNGLMKKIDETNVDMLIDPVPICDLMALVLPVLTIRISVGNMER
SCGLPAPLSYVPVPMGTGLTDMTEFLERVKNSMLSVLFHFIQDYDFHWEFEYSKAL
GRPTLCETVGAKEIWLIRTYWDFEPPOYPQNFVGLHCKPAKALPKWENFVQS
SEGDGIVFSLGSFNQNTVEERKANTIASALAOIPKVLWRYGKPKPSLGNANRLIDW
IPONDILHGPRTKAPTTHGGMNGIYEALYHGVPMVGVPIFGDOLDINAHMKAKAAVE
INFKYMSDELLARLRTVITDSYKENAMRLSRIRHDDQPVKPLDRAVEVFEFVWRHKG
AKHLRSAAHDLTFWQHSIDVIGFLACVATAIFLFTKFLFCQKFNKRKIEKRE"

BASE COUNT 919 a 463 c 496 g 908 t

ORIGIN

Query Match 88.3%; Score 1444.4; DB 9; Length 2786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	191	ATTGACTCACTCAAGCCCTTCGTTAATTGACTACAGGAGCCTTCTGCAATTGAATTTGA	250
Db	1	ATTGACTCACTCAAGCCCTTCGTTAATTGACTACAGGAGCCTTCTGCAATTGAATTTGA	60
QY	251	GCTGTCCATATGCCACAGGACAGAACAGAGAAATGAAATATTTGTTGCTACGCTCT	310
Db	61	GCTGTCCATATGCCACAGGACAGAACAGAGAAATGAAATATTTGTTGCTACGCTCT	120
QY	311	GAATGCTTGCAGGCTTATCAACCTGGCAATCACTGTTATAAAATTAATGATTTTGT	370
Db	121	GAATGCTTGCAGGCTTATCAACCTGGCAATCACTGTTATAAAATTAATGATTTTGT	180
QY	371	TGAATAAGAGAACTTTAAATAATGATGTGAGAGCTTTATCTACAAATCAGAGCTTAT	430
Db	181	TGAATAAGAGAACTTTAAATAATGATGTGAGAGCTTTATCTACAAATCAGAGCTTAT	240
QY	431	GAAGAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCTGTGATTCCTGTGG	490
Db	241	GAAGAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCTGTGATTCCTGTGG	300
QY	491	AGACCTGATGGCTCAGTTGCTTGCAGTCCCTTTTGCTGCTCACACTTGAATTTCTGAG	550
Db	301	AGACCTGATGGCTCAGTTGCTTGCAGTCCCTTTTGCTGCTCACACTTGAATTTCTGAG	360
QY	551	AGGCAATATGGAGGAAAGCTGTGGGAACTTCCAGCTCCACTTTCTATGCTACCTGTGC	610
Db	361	AGGCAATATGGAGGAAAGCTGTGGGAACTTCCAGCTCCACTTTCTATGCTACCTGTGC	420
QY	611	TATGACAGGACTACACAGAGATGACCTTTCTGGAAAGATGAAAAATTCATGCTTTC	670
Db	421	TATGACAGGACTACACAGAGATGACCTTTCTGGAAAGATGAAAAATTCATGCTTTC	480
QY	671	AGTTTTGTTCCACTTCTGGATTACGAGTATACGACTATCATTTTGGGAGAGTTTTATAG	730
Db	481	AGTTTTGTTCCACTTCTGGATTACGAGTATACGACTATCATTTTGGGAGAGTTTTATAG	540

QY	731	TAAGCATTTAGGAGGCCACCTACATTTATGTGAGACTGTGGGAAAAGCTGAGATATGGCT	790
Db	541	TAAGCATTTAGGAGGCCACCTACATTTATGTGAGACTGTGGGAAAAGCTGAGATATGGCT	600
QY	791	AATACGAACATATTTGGGATTTTGAATTTCTCAACCATACCAACCTTAACCTTTGAGTTGT	850
Db	601	AATACGAACATATTTGGGATTTTGAATTTCTCAACCATACCAACCTTAACCTTTGAGTTGT	660
QY	851	TGGAGGATTTGCACTGTAACCTGCCAAAGCTTTGCTTAAGGAAATGAAAAATTTGTCCA	910
Db	661	TGGAGGATTTGCACTGTAACCTGCCAAAGCTTTGCTTAAGGAAATGAAAAATTTGTCCA	720
QY	911	GAGTTTCAGGGGAGAGATGATTTGTTGTTCTCTCTGGGGTCACTGTTTCAAAATGTTAC	970
Db	721	GAGTTTCAGGGGAGAGATGATTTGTTGTTCTCTCTGGGGTCACTGTTTCAAAATGTTAC	780
QY	971	AGAGAAAAGGCTAATATCATTTGCTTCAGCCCTTTCGCCAGATCCACAGAGGTGTTATG	1030
Db	781	AGAGAAAAGGCTAATATCATTTGCTTCAGCCCTTTCGCCAGATCCACAGAGGTGTTATG	840
QY	1031	GAGGTACAAAGGAAAGAAACCATCCACATTTAGGAGCCAACTACTCGGCTGATGATGGAT	1090
Db	841	GAGGTACAAAGGAAAGAAACCATCCACATTTAGGAGCCAACTACTCGGCTGATGATGGAT	900
QY	1091	ACCCAGAAATGATCTTCTTGGTCAATCCAAACCAAGCTTTTATCACTCATGTTGGAAT	1150
Db	901	ACCCAGAAATGATCTTCTTGGTCAATCCAAACCAAGCTTTTATCACTCATGTTGGAAT	960
QY	1151	GAATGGGATCTATCAAGCTATTTACCATGGGGTCCCTTATGGTGGGAGTTCCTCATATTTGG	1210
Db	961	GAATGGGATCTATCAAGCTATTTACCATGGGGTCCCTTATGGTGGGAGTTCCTCATATTTGG	1020
QY	1211	TGATCAGCTTGTATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAACTT	1270
Db	1021	TGATCAGCTTGTATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAACTT	1080
QY	1271	CAAACTATCACAGGAGAGATTTACTGGGGCTTTGAGAACAGTCACTTACCGATTCTTC	1330
Db	1081	CAAACTATCACAGGAGAGATTTACTGGGGCTTTGAGAACAGTCACTTACCGATTCTTC	1140
QY	1331	TTATAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTGTAAAGCCCT	1390
Db	1141	TTATAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTGTAAAGCCCT	1200
QY	1391	AGATCAGAGAGTCTTCTGGATCGAGTTTCTATGCCCAACAAAGGAGCCACACCTGGG	1450
Db	1201	AGATCAGAGAGTCTTCTGGATCGAGTTTCTATGCCCAACAAAGGAGCCACACCTGGG	1260
QY	1451	ATCAGCTGCCATGACCTCACCTGTTCCAGCAGCTACTCTATAGATGTGATGGGTTCT	1510
Db	1261	ATCAGCTGCCATGACCTCACCTGTTCCAGCAGCTACTCTATAGATGTGATGGGTTCT	1320
QY	1511	GCTGACCTGTGGCACTGCTATATCTTGTTCACAAAATGTTTTTTTATTTCTCTGTCA	1570
Db	1321	GCTGACCTGTGGCACTGCTATATCTTGTTCACAAAATGTTTTTTTATTTCTCTGTCA	1380
QY	1571	AAAAATTAATAAATAG	1630
Db	1381	AAAAATTAATAAATAG	1440
QY	1631	GACCTG 1636	
Db	1441	GACCTG 1446	

RESULT 5
AF175221
LOCUS
DEFINITION
Cavia porcellus UDP glucuronosyltransferase UGT2A3 mRNA, complete cds.
AF175221
ACCESSION
AF175221.1 GI:5802603
VERSION

AF175221 2547 bp mRNA linear ROD 18-MAY-2001
Cavia porcellus UDP glucuronosyltransferase UGT2A3 mRNA, complete
AF175221
AF175221.1 GI:5802603

QY 1506 TTCCTGCTGCTCTGTGGCACTGCTATATCTTGTTCACAAAATGTTTTTTATTTCC 1565
Db 1508 TTCCTGCTGCTGCTGTAGCAAGTCTATATGCTGTGTCACAAAATGCTGCTTATTTCC 1567
QY 1566 TGTCAAAATTTTATAAACTAGAAAGATAGAAAGAGGCAATAGATCTTCCAAAT 1623
Db 1568 TTTCAAACTTTTATTAAGATAGAAAGAGATAGAAAGAGATAGATCTTCAAAAT 1625

RESULT 6
LOCUS AX040087 1650 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 10 from Patent WO063351.
ACCESSION AX040087
VERSION AX040087.1 GI:11230049
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1650)
AUTHORS Lal, P., Yue, H., Tang, Y. T., Hillman, J. L., Baughn, M. R. and Yang, J.
TITLE Carbohydrate-modifying enzymes
JOURNAL Patent: WO 063351-A 10 26-OCT-2000;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 2912330CBI"
BASE COUNT 489 a 330 c 354 g 477 t
ORIGIN

Query Match 59.1%; Score 967.4; DB 6; Length 1650;
Best Local Similarity 75.7%; Pred. No. 1.4e-239;
Matches 1228; Conservative 0; Mismatches 386; Indels 9; Gaps 2;

QY 23 CATCATGAGCTGTGACAGTGTGCTTGTGATTTCTGCTCTGAGCTCTCTCTGT 79
Db 28 CATCAGGATGCTGTGAAGTGTGCTGCTGCTCTCTGATACAGCTGAGCTGTACTT 87
QY 80 TGGCTGTGGATTTCTGTGGAAAGTCTGTGCTGTGCTGCTGCTGCTGCTGCTGCT 139
Db 88 TAGCTCTGGAGTGTGCAAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147
QY 140 TGTCAAGTCAATCTAGAGAGCTCATAGTGAAGGCTGAGTGAACAGTATTGACTCA 199
Db 148 TATAAGACAACTCTGGATGAATCTGTCAGAGAGGTCATGAGTGTGCTGCTGCTGCT 207
QY 200 CTCAAAGCTCTGCTTAATGACTACAGAGGCTTCTGCAATGAAATTTGAGTGTGCTCA 259
Db 208 TTCAGCTTCCATTTCTTCTGCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
QY 260 TATCCACAGACAGACAGAGAAATGAATATTGTTGACTGCTGCTGCTGCTGCTGCTG 313
Db 268 TGTATCTTTAACTAAACTGAGTTGAGGATATTATCAAGCAGCTGTTGAAGAGATGGGC 327
QY 314 TGTCTTCCAGGCTTATCACTTGGCAATCAGTTTATAAATTAATTAATTTTGTGTA 373
Db 328 AGAATTTCCAAAGACACATTTTGGTCATATTTTCAAGTACAAGTAACAAGTATCATG 387
QY 374 AATAAGAGGAACCTTTAAATATGATGTGTGAGAGCTTTATCTACAATCAGACGCTTAT 433
Db 388 ATTATATGACATACTTAGAAAGTCTGTAAAGGATATAGTTTCANATAAGAAATATGAA 447
QY 434 GAAGCTACAGGAACCACTACAGTATATGCTTATAGACCTGTGATTCCTGCTGGAGA 493
Db 448 GAAGCTACAGGAGTCAAGATTTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507
QY 494 CCTGATGGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
Db 508 GCTGCTGGCGGATTTAAATATACCTTTGTGTACAGCCTCTCTCTCTCTCTCTCTCT 567

QY 554 CAATATGAGCGAAGCTGTGGGAAACTTCCAGCTCCACTTCCCTATGTACCTGTGCCTAT 613
Db 568 CCCAATTGAAAAGCATAGTGGAGGACTTCTGTCTCCCTCTCTCTATGTATGCTGTGTAT 627
QY 614 GACAGACTTAACAGACAGAACTGACCTTCTGGAAGAGTAAAAAATTCAAATGCTTTCAGT 673
Db 628 GTCAGAACTAAGTGACCAAAATGACCTTCATAGAGAGGTAATAAATATGATCTATGTCT 687
QY 674 TTTGTTCCACTTCTGGATTTCAGGATTCAGCACTATCATTTTGGGAAGAGTATTTATAGTAA 733
Db 688 TTATTTGAATTTTGGTTCCAAATATTGACATGAAGAAGTGGGATCAGTTCACAGTGA 747
QY 734 GGCATTAGGAAGGCCACTACATATATGTGAGACTGTGGGAAAGCTGAGATATGCTTAAT 793
Db 748 AGTCTAGGAAGGCCACTACGTTATCTGAGACAATGGCAAAAGCTGACATATGCTTAT 807
QY 794 ACGAACATATTGGGATTTTGAATTTCTCAACCATAGCAACCTAACCTTTCAGTTCGTTGG 853
Db 808 TCGAAACTACTGGGATTTTCAATTTCTCACCACCTTACCAAAATGTTGAGTTCGTTGG 867
QY 854 AGGATTCGACTGTAAACCTGCCAAAGCTTTCCTTAAGGAAATGGAATAATTTGTCAGAG 913
Db 868 AGGACTCCACTGCAAACTGCCAAACCCCTACCGAAGGAAATGGAAGATTTGTCAGAG 927
QY 914 TTCAGGGAAGATGATTTGTGTGTTTCTCTGCGGTCACTGTTTCAAAATGTTACAGA 973
Db 928 CTCTGGAGAAATGTTGTTGTTCTCTGCGGTGATGTTGAGTAACACGTCAGA 987
QY 974 AGAAAGGCTTAATATCATCTGCTGAGCCCTTGGCCAGATCCACAGAGGTGTTATGGAG 1033
Db 988 AGAAAGGCTTAATGTTGTTGTTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1047
QY 1034 GTCAAGGAAAAAACCATCCACATAGGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTG 1093
Db 1048 ATTTGATGGGAATAACACGATACTTAGGCTCAATCTGCTGCTGCTGCTGCTGCTGCTG 1107
QY 1094 CCAGATGATCTTCTTGGTGTATCCCAAAACCAGCTTTTATCCTCATCTGCTGCTGCTG 1153
Db 1108 CCAGATGATCTTCTTGGTGTATCCCAAAACCAGCTTTTATCCTCATCTGCTGCTGCTG 1167
QY 1154 TGGGATCTATGAAGCTATTTACCATGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1213
Db 1168 TGGGATCTATGAAGCTATTTACATGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
QY 1214 TCAGCTTGATAACATAGCTCAGTGAAGGCCAAAGGAGCAGCTGTAGAAATAAATTCCTCA 1273
Db 1228 TCAAGCTTGATAACATAGCTCAGTGAAGGCCAAAGGAGCAGCTGTAGAAATAAATTCCTCA 1287
QY 1274 AACTATGACAAGGAGATTTTACTGAGGCTTTTGAAGAGCTATTCAGTTCCTCTTA 1333
Db 1288 AACTATGACAAGGAGATTTTACTGAGGCTTTTGAAGAGCTATTCAGTTCCTCTTA 1347
QY 1334 TAAAGAGATGCTATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1393
Db 1348 TAAAGAGATGCTATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 1407
QY 1394 TCGAGCAGCTTCTGCTGATCGAGTTTGTGATCGGCTTCAAGAGGCTTCAAGAGGCTTCA 1453
Db 1408 TCGAGCAGCTTCTGCTGATCGAGTTTGTGATCGGCTTCAAGAGGCTTCAAGAGGCTTCA 1467
QY 1454 AGCTGCCCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1513
Db 1468 AGCTGCCCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527
QY 1514 GACCTGTGTGGCACTGCTATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1573
Db 1528 GACCTGTGTGGCACTGCTATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1587
QY 1574 ATTTAAATAAACTAGAAAGATAGAAAGGAGATAGATCTTCCCAATTTCAAGAAAGAC 1633
Db 1588 ATTTAAATAAACTAGAAAGATAGAAAGGAGATAGATCTTCCCAATTTCAAGAAAGAC 1647


```
Db 962 AGAGAAAGCCAACTCATTTGCCTCTGTCTTGGCCAGATTCCCCAGAAAGGTTTTGTGGAG 1021
Qy 1034 GTACAAGAGNAAAAACCAATCCACATTAGAGCCAACTACTCGGCTGTATGATTTGGATACC 1093
Db 1022 ATACTCAGGCAAGAACCCAGCCACATTAGGATCCAACTACTCGGCTTTTAAATGGATGCC 1081
Qy 1094 CCAGAATGATCTCTTGGTTCATCCCAAAACCAAGCTTTTATCACTCATCTGGTGGGAATGAA 1153
Db 1082 CCAGAATGATCTCTTGGACATCTTAAACCAAGCTTTTATCACTCATCTGGTGGGAATGAA 1141
Qy 1154 TGGGATCTATGAAGCTATTATACATGGGCTCCCTATGTTGGGAGTTCCTCATATTTGGTGA 1213
Db 1142 CGGGATTTATGAAGCCATTATCAGTGGGCTCCCTATGTTGGGCGTTCCCATGTTAGGGA 1201
Qy 1214 TCAGCTTGATAACATAGCTCACATGAAGCCCAAGGAGGAGCTGTAGAAATAAATCTCAA 1273
Db 1202 TCAGCCTCACATCTGCTCATGAGGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1261
Qy 1274 AACTATGACAGCGAAGATTACTAGGCGCTTTGAGAACAGCTCATACCGGATTCCTCTTA 1333
Db 1262 TACAATGACGAGCAGACAGATTACTCAGTGTCTGAGGCGAGTCATATGAGCCCTCTTA 1321
Qy 1334 TAAAGAGATGCTATGAGATATCAAGATTTCACCATGATCAACCTGTAAAGCCCTAGA 1393
Db 1322 TAAAGAGATGCCATGCGGTTATCAAGAAATCCACATGATCAGCGAGTGAAGCCCTGGA 1381
Qy 1394 TCGAGCAGCTCTCTGGATCGAGTTTGTCTATCGGCGCCACAAAGGAGGAGGAGGAGGAGGAG 1453
Db 1382 CCGAGCAGCTCTCTGGATGAGTTGTCATGCGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1441
Qy 1454 AGCTGCCCATGACCTCAGCTGGTTCAGCAGCTACTCTATAGATGATGATGGGTTCCCTGCT 1513
Db 1442 GCGAGCCCATGACCTCAGCTGGTTCAGTACCACTCCCTAGATGATGATGGGTTCCCTGCT 1501
Qy 1514 GACCTGTGGCACTGCTATATCTTGTTCACAAATGTTTTTATTTTCTCTGTCGAAAA 1573
Db 1502 GTTGTGTGCTTACTCTGACATTCATCACTCACTAAATTTTGTGTGTGTCGAAAA 1561
Qy 1574 ATTAAATAAAGTGAAGATAGAAAAGAGGGAATAGA 1611
Db 1562 ACTTATATATGA-AAGAAAGTGAAGAAATAGGGAACAGA 1598

RESULT 8
LOCUS HSA6054 1766 bp mRNA linear PRI 16-JUL-1999
DEFINITION Homo sapiens mRNA for UDP glucuronosyltransferase.
ACCESSION AJ006054
VERSION UDP glucuronosyltransferase; GI:4753765
KEYWORDS UDP glucuronosyltransferase;ugt2A1 gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1766)
Jedlitschky,G., Cassidy,A.J., Sales,M., Pratt,N. and Burchell,B.
Cloning and characterization of a novel human olfactory
UDP-glucuronosyltransferase
Biochem. J. 340 (Pt 3), 837-843 (1999)
92889328
MEDLINE 10359671
PUBMED
REFERENCE
1 (bases 1 to 1766)
Cassidy,A.J.
Direct Submission
Submitted (07-MAY-1998) Cassidy A.J., Molecular and Cellular
Pathology, University of Dundee, University Dept. of Molecular and
Cellular pathology, Ninewells Hospital, Dundee, DD1 9SY, SCOTLAND
Location/Qualifiers
1. .1766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="olfactory"
64. .1647

FEATURES
source
```

CDS

```
/gene="ugt2A1"
64. .1647
/ene="ugt2A1"
/EC_number="2.4.1.17"
/codon_start=1
/product="UDP glucuronosyltransferase"
/protein_id="CAB41974.1"
/db_xref="GI:4753766"
/translation="MLNNLLLSLQISLIGTTLGGNVLIMPMEGSHLWNLKIIIDELI
KERNVIVLVASGALFIPITSPNSLTFEIKYKPEGKERIEGVIKDFVSTWLENRSPS
TIMPYQEMAKVINDKDFHMSVOEICDGVLMKQMLAKLKFVLDVSDVPFGDIVA
LKLGIPMSYLSRSPASTVBKHCQKVPYPSYPAVLSELTDDMSFTDRINRISYHL
QDYMFETLWKSWSYSYKALGRPTTLCETMGKAEIWLIRTYWDFEPFPLPPEFVG
GLHCKPAKPLPKMEBEFIOSSKNGVVFSLGSMVKNLTEEKANLIALSALQAPKQVL
WRYKGGKPAITLGNNTOLFDAIPQNDLLGHFKAFITHGHTNGIYEAIYHGVPMVGP
MFADOPNIAHMKAKGAAREVENLNTMSVDLLSALRTVINEPSYKENAMLSRIHDDQ
PVKPLDRAVFWIEFVMRHKGAKHLRVAHDLTWFOYHSLDVIIGLLVCVTTAFLVIO
CCULFSCKQFKIGKKKRE"
BASE COUNT 542 a 359 c 363 g 502 t
ORIGIN
Query Match 50.2% Score 821.2; DB 9; Length 1766;
Best Local Similarity 71.9%; Pred. No. 1e-201;
Matches 1107; Conservative 0; Mismatches 418; Indels 15; Gaps 2;
Qy 94 GTGGAAAGTCCTGGTGTGGCCCTGTGACATGAGCCATTGGCTTAATGTCAAGGTGCTATTC 153
Db 122 GTGGGAATGTTTGGATTGGCCAATGGAAGTAGTCATTGGCTAAATGTTAAGATAATTA 181
Qy 154 TAGAAGAGCTCATAGTGAGAGGCCATGAGTAACAGTATTGACTCACTCAAGACCTTCGT 213
Db 182 TAGTAGAGCTCATATAAAGAGGAGCAATGTAAGTGTCTCTAGTGTGCTGTGCTGCTT 241
Qy 214 TAATTGACTACAGGAAGACCTTCTGCAATGAAATTTGAGGTGGTCCATATATGCCACAGACA 273
Db 242 TCATCACACCAACCTCTAACCCTCTGACATTTGAATATATAGGTGCCCTTTGGCA 301
Qy 274 GAACAGAGAAATGAATATTTGTTGACCTAGCTCTGAATGCTTTGCCAGGCTTATCAA 333
Db 302 AAGAAAGATAGAAGGAGTAATTAAGGACTTCGTTTCGACATGGCTGGAAATAGACCAT 361
Qy 334 CTGGCAATCAGTTATAAAATTAATGATTTTTCCTGAAATAGAGGAACCTTTAAAAA 393
Db 362 CTCCTTCAACCATTTGGAGATTCTATCAGAGATGCCAAAGTAATCAAGAGACTTCCACA 421
Qy 394 TGATG-----TGTGAGAGCTTTTATCAATCAGACGCTTATGAAGAAAGCTAC 441
Db 422 TGGTGTCTCAGGAGATCTGTGATGGCGTCTTTAAAAACCAACAGCTGATGGCAAGCTAA 481
Qy 442 AGGAAACCAACTACGATGTAATGCTTATAGACCTGTGATTCCTCTGGAGACCTGTATGG 501
Db 482 AGAAGAGCAAGTTTGAAGTCTTGGTGTCTGATCCAGTATTCCTTGTGGCGATATAGTAG 541
Qy 502 CTGAGTGTGCTCAGTCCCTTTTGTGCTCACACTTTAGAATTTCTCTAGGAGCAATATGG 561
Db 542 CTTTAAACCTTGAATTCATTTATGTAATCTCTGAGGTTTTCTCCAGCCTCAACAGTGG 601
Qy 562 ACCGAGAGCTGTGGAAACTTCCAGCTCCACTTCCATATGATGCTGTGCTGTGCTATGACAG 621
Db 602 AAAAGCACTGTGGGAAGGTACCATACCCTCTCTCTATGTTCTCTCTGTTTATCAGAAC 661
Qy 622 TAACAGACAGATGACCTTTCTGGAAGAGTAAAAAATTCATGCTTCAGTTTGTGTTCC 681
Db 662 TCACCGACCAATGCTTCTACTGACAGAAATAAGAAAT---TCATCTCTACCACTTAC 718
Qy 682 ACTTCTGAGTTCAGGATTCAGCACTATCATTTTGGGAAGAGTTTTATAGTAAGCACTAG 741
Db 719 AGGACTACATGTTGAAACTCTTTGGAATCATGSGATTTCATATAGTAAGCACTTAG 778
Qy 742 GAAGCCCACTACATTATGTCAGACTGTGGGAAGAGTGGAGATATGGCTAATACGAACAT 801
Db 779 GAAGACCACCTACGTTATGTGAGACTATGGGGAAGAGCTGAAATTTGGTTAATCCGAACAT 838
```

Qy	802	ATTGGGATTTTGAATTTCTCAACCATACCAACCTAACTTTGAGTTTGTGGAGGATTCG	861
Db	839	ATTGGGATTTTGAATTTCTCGTCCATACATTACCTAATTTTGAATTTTGTGGAGGATTCG	898
Qy	862	ACTGTAACACTGCCAAAGCTTTCGCTAAGCAATGGAAATTTTGTCCAGATTCAGGGG	921
Db	899	ACTGCAAACTGCCAAACCTTTTACCTAAGGAATGGGAAGATTTATCCAGAGCTCAGGTA	958
Qy	922	AAGATGGTATGTGGTGTCTCTCTGGGGTCACTGTTTCAAAATGTTTACAGAAGAAAAGG	981
Db	959	AAAATGCTGTGGTGTCTCTCGGATCAATGGTCNAAAACCTTACAGAAGAAAAGG	1018
Qy	982	CTAATATCATTTGCTTCAGCCCTTGCCCAGATCCCAAGAGGTGTTATGGAGGTACAAG	1041
Db	1019	CCAATCTTATTGGCTACAGCCCTTGCCCAGATTTCCACAGAAGGTTTTATGGAGATACAAG	1078
Qy	1042	GAIAAAACCATCCACATTAGGAGCCAACTACTCGGCTGATGATGGATACCCCAAGATG	1101
Db	1079	GAIAAANACGACCAATTAGGAACAACTACTCAGCTCTTTGATTTGGATACCCCAAGATG	1138
Qy	1102	ATCTTCTTGTGTCATCCCAAAACCAAGCTTTTTATCACTCATGTGTGGAATGAATGGGATCT	1161
Db	1139	ATCTTCTTGGACATCCCAAAACCAAGCTTTTTATCACTCATGTGTGGAACCTAATGGGATCT	1198
Qy	1162	ATGAGCTATTACCATGGGGTCCCTATGTGGGAGTTCCCATATTTTGGTGTATCAGCTTG	1221
Db	1199	ACGAAGCTATTATTACCAAGGATCCCTATGTGGGAGTTCCCATATTTTGTCTGATCAGGCTG	1258
Qy	1222	ATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCCAAACATAGA	1281
Db	1259	ATAACATTGCTCACATGAAGGCCAAAGGAGCAGCTGTGGAAGTCAACCTAAACACATGA	1318
Qy	1282	CAGCGAAGATTTACTGAGGGCTTTGAGAACAGTCATTACCGATTCCTCTTATAAGAGA	1341
Db	1319	CAAGTGTGGATTTTGTCTTAGCGCTTTGAGAACAGTCATTAAATGAACCTTCTTATAAGAGA	1378
Qy	1342	ATGCTATGAGATTATCAAGAATTCACCATCATCAACCTGTAAAGCCCTAGATCGAGCAG	1401
Db	1379	ATGCTATGAGTTATCAAGAAATTCACCAATCATCAACCTGTAAAGCCCTGGATCGAGCAG	1438
Qy	1402	TCCTTCTGGATCGAGTTTGTCTATGGGCCAAAGAGGACCAAGCACCCTGCGATCAGCTGCC	1461
Db	1439	TCCTTCTGGATCGAGTTTGTCTATGGGCCAAAGAGGACCAAGCACCCTTCGGGTTTCAGGCC	1498
Qy	1462	ATGACCTCACCCTGGTTCACGACTACTCTATAGATGTGATTTGGGTTCTCTGCTGACCTGTG	1521
Db	1499	ATGACCTCACCCTGGTTCACGACTACTCTATAGATGTGATTTGGGTTCTCTGCTGCTGTG	1558
Qy	1522	TGGCAACTGCTATATTCTTCTCACAAAATGTTTTTTTATTTCTCTGTCAAAAATTTAATA	1581
Db	1559	TGACAGGCTATATTTTGGTCTACAAATGTTGTTTGTGTTTCTCTGTCAAAAATTTGGTA	1618
Qy	1582	AAACTAGAAAGATAGAAAGAGGGAATAGATCTTTTCCAAA	1621
Db	1619	AGATAGAAAGAGAAAGAAAGAAAGATAGCTCAAGAAAAA	1658

RESULT	9
AX410646	
LOCUS	2090 bp DNA linear PAT 14-JUN-2002
DEFINITION	Sequence 3293 from Patent WO0229103.
AX410646	
ACCESSION	AX410646.1 GI:21443351
VERSION	
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1
REFERENCE	Alvares C., Horne D., Peres-da-silva S. and Vockley J.G. Gene expression profiles in liver cancer Patent: WO 0229103-A 3293 11-APR-2002;
AUTHORS	
TITLE	
JOURNAL	GENE LOGIC INC (US)

Db 496 GAGCTACTGGCTGAACATATTTAAACATACCCCTTTCTGTACAGTCTTCGATCTCTCTGTGGC 555
QY 552 GGCATATGAGCGAGAGCTGTGGAAACTTCCAGCTCCACTTTCCTATGTACCTGTGGCT 611
Db 556 TACACATTTGAGAAGATGGTGGAGGATTTCTGTCCCTCTCTCTATGTACCTGTGTT 615
QY 612 ATGACAGGACTAACACAGACAGATGACCTTCTTGGAAAGAGTAAATAATCAATGCTTTCA 671
Db 616 ATGTCAGAATTAAGTGATCAATGATTTTCATGGAGAGGATAAAAAATATGATACATATG 675
QY 672 GTTTTGTCCACTTCTGTGATCAGGATTAACAGTATCAATTTTGGGAAGAGTTTATAGT 731
Db 676 CTTTATTTGATTTTGTGTTTCAAAATTTATGATCTGAAGAAGTGGGACCACTTTTATAGT 735
QY 732 AAGCATTTAGGAAGCCCACTACATATGAGAGTGTGGGAAGAGTGAATATGGCTA 791
Db 736 GAAGTTCTAGGAAGACCCACATATTTGAGACAATGGGAAGCTGGAATGTGGCTC 795
QY 792 ATACGAACATATTTGGGATTTTGAATTTTCTCAACCATACCAACCTTACCTTTGAGTTTGT 851
Db 796 ATTCGAACCTATTTGGGATTTTGAATTTTCTCGCCCATTTCTTACCAAAATGTTGATTTGTT 855
QY 852 GGAGGATTCGACTGTAACCTGCGCAAGCTTTGCTTAAGGAATGGAATAATTTGTCCAG 911
Db 856 GGAGGACTTCACTGTAAACCAAGCCCAACCCCTGCTTAAGGAATGGAAGATTTGTGCGAG 915
QY 912 AGTTCAAGGGAAGATGATGTTGTTGTTCTCTGCGGTGCTGTTTCAAAATGTTTACA 971
Db 916 AGCTCTGGAAGAAATGATGTTGTTGTTCTCTGCGGTGCTGTTTCAAAATGTTTACA 975
QY 972 GAAGAAAGAGGTAATATCATTTGCTTACGCCCTTGGCCAGATCCCAAGAGAGTGTATGG 1031
Db 976 GAAGAAAGTGCCACATGATTTGCTATGCTAGCCCTTGGCCAGATCCCAAGAGAGTGTATGG 1035
QY 1032 AGGTACAAGGAAGAAACCATCCATAGAGCCATATCTCGGCTGTATGATGGATA 1091
Db 1036 AGATTGTAGCAAGAGCCCAATACATTTAGTTTCCAACTACTCGACTGTACAAGTGGTA 1095
QY 1092 CCCAGATGATCTTCTTGGTTCATCCCAAAACCAAGCTTTTATCACTCATGGTGAATG 1151
Db 1096 CCCAGATGACCTTCTTGGTTCATCCCAAAACCAAGCTTTTATCACTCATGGTGAATG 1155
QY 1152 AATGGATCTATGAGCTATTTACATGGGTGCTTATGTTGGAGATTTCCCATATTTGGT 1211
Db 1156 AATGGCATCTATGAGGCTATCTACCATGGATCCCTATGTTGGGCTATCCCTTGTGTCG 1215
QY 1212 GATCAGCTGTATACATAGCTACATGAGGCGCAAGAGAGAGCTGTAGATAAATCAATC 1271
Db 1216 GATCAACATGATAACATTTGCTACATGAAGCCCAAGGAGAGGAGCCCTCAGTGTGGACATC 1275
QY 1272 AAACTATGACAAGCGAAGATTTACTGAGGGCTTTGAGAAGAGTCAATACCGATTCCTCT 1331
Db 1276 AGGACCATGTCAGTAGAGATTTGCTCAATGATGAGTCACTATTAAGTACCCCTGTC 1335
QY 1332 TATAAGAGATGATGAGATTTATCAAGATTTACCATGATCAACCTGTAAAGCCCTCA 1391
Db 1336 TATAAGAGATGATGAGATTTATCAAGATTTATCAATGATCAACCAATGAAGCCCTG 1395
QY 1392 GATCAGAGCTTCTGATGAGTGTGCTATGCGCCCAAGAGAGGAGCCACCTGCGA 1451
Db 1396 GATCAGAGAGCTTCTGATGAGTGTGCTATGCGCCCAAGAGAGGAGCCACCTGCGA 1455
QY 1452 TCAGTGCCTATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGATGGTTCCTG 1511
Db 1456 GTCGAGCTCACAACTCACTGGATCCAGTACCACTCTTTGGATGTGATAGCATTCCTG 1515
QY 1512 CTGACCTGTGGCAACTGCTATATCTTGTGTCAAAATGTTTTTATTTTCCCTGTCAA 1571
Db 1516 CTGGCTCGTGGCAACTGCTATATTTATCATCAAAATTTTGGCTGTTTGTGTTCCGA 1575
QY 1572 AATTTTAACTAGAAAGATAGAAAGAGGAGGATAGATCTTCCAAA 1621
Db 1576 AAGCTTGCCAAACAGGAAGAAAGAAAGAGATTTAGTTATATCAAAA 1625

RESULT 11
AF180322 2114 bp mRNA linear PRI 14-SEP-1999
LOCUS Homo sapiens UDP-glucuronosyltransferase 2B15 (UGT2B15) mRNA,
DEFINITION UGT2B15-Y85 allele, complete cds.
ACCESSION AF180322
VERSION AF180322.1 GI:5881245
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2114)
AUTHORS Levesque,E., Beaulieu,M., Green,M.D., Tephly,T.R., Belanger,A. and Hum,D.W.
TITLE Isolation and characterization of UGT2B15(Y85): a
UDP-glucuronosyltransferase encoded by a polymorphic gene
JOURNAL Pharmacogenetics 7 (4), 317-325 (1997)
MEDLINE 97439504
PUBMED 9295060
AUTHORS 2 (bases 1 to 2114)
Levesque,E., Beaulieu,M., Green,M.D., Tephly,T.R., Belanger,A. and Hum,D.W.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1999) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Ste-Foy, Quebec G1V 4G2, Canada
FEATURES
source
1..2114
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q13"
1..2114
/gene="UGT2B15"
/allele="Y85"
39..1631
/gene="UGT2B15"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B15"
/protein_id="AAD5093.1"
/db_xref="GI:5881246"
/translation="MSLKWTSVFLLIQLSYFSSGCGKVLVWPTXSHWINMKTILE
ELVORGHEVTLTSSAFLVNAKSSAIKLEVYPTSLTKNYLEDLILDLRWYGV
KNTWVSFQLOELCWYDYDNKLDKAVLNKMKLOESKFDVLIADALNPGEL
LAELENPFLYSLRESYGTPEKNGGFLFPSPVPMVMSLSQIMFMRKNMIM
LYDFWQIYDLKKWDQYSEVLGRPTTLFETMKAEMMLIRTWDFEPPPLPND
FVGLHCKPAKPLPKMEEFVQSSGNGIVFVSLGSMISNSESANNIASALAOIPO
KVLRFQDKPNTLGSNTRLYKLPQNDLLGHPKTKAFITHTGNGIYEATYHGPIV
GIPLFADQHDNIAHMKAKAALSVDIRTMSSRDLLNALKSVINDPVYKENVKLSRIH
HDQPMKPLDRAVFWIEFVNRHKAKHLRVAHNLTWIQYHSLDVIAPLLACVATVFI
ITKCLCFRKLAKTKKKRD"

BASE COUNT 670 a 386 c 435 g 623 t

ORIGIN
Query Match 47.6%; Score 778.8; DB 9; Length 2114;
Best Local Similarity 68.7%; Pred. No. 1e-190;
Matches 1106; Conservative 0; Mismatches 492; Indels 12; Gaps 2;
QY 24 ATCATGAGGTCTGCAACATCAGCTTTGGTATTTCTTCGCTCCGACCTCTCTCTGT 80
Db 33 ACCAGGATGCTCTCTGAAATGGACGTCAGTCTTCTGCTGATACAGCTCAGTTTACTTT 92
QY 81 GCCTGTGATTTCTGTGGCAAGTCTCTGCTGTGGCCCTCTGACATGAGCATTGGCTTAAT 140
Db 93 AGCTCTGGAAGCTGTGGAAGAGGTGCTAGTGTGGCCCAAGATACAGCACTGGATAAAT 152
QY 141 GTCAGGCTCATTTAGAGAGCTCATAGTGTAGAGGCCATGAGGTAAACAGTATTGACTCAC 200
Db 153 ATGAAGACAATCCTGGAAGAGCTTCTTCAGAGGGGTGATGAGTGACTGTGTTGACATCT 212
QY 201 TCAAGACCTTCGTTAAATTTGACTACAGGAAGCCTTCTGCAATTTGAATTTGAGGTGCTCAT 260

Db	406	GAATATTTCTGACATATAATATAAAGCTGCTGGAAGATGACGATTTTGAACAAGAACCTTATG	465
Qy	432	AAGAAGCTACAGAAACCAACTACGATGTAAATGCTTATAGACCCCTGTGATTCCTCTGGGA	491
Db	466	AGAAACTACAGAGTCAAAATTTGATGTCTTCTGGCAGATGCCGTTAATCCCTGTGGT	525
Qy	492	GACCTGATGCTGAGTTGCTTTGTCAGTCCCTTTTGTGCTCACACTTAGAATTTCTGTAGGA	551
Db	526	GAGCTGTGGCTGAACCTACTTAACATACCCCTTTCTGTACAGTCCCGCTTCTCTGTGGC	585
Qy	552	GGCAATATGAGCGAAGCTGTGGAAACTTCCAGCTCCACTTTTCTATGTACCTGTGCGCT	611
Db	586	TACACAGTTGAGAAGAAATGTGTGAGGAGTTTCTGTCCCTCTTCTATGTACCTGTGT	645
Qy	612	ATGACAGACTTAACAGACAGACAGTACCTTCTTGAAAGAGTAAAAAATCAATGCTTTCA	671
Db	646	ATGTCAGAAATTAAGTGAATCAATGATTTTCTATGGAGAGGATAAAAAATATGATATATG	705
Qy	672	GTTTTGTTCACCTTCGTGATTCAGGATTAACGACTATCATTTTGGGAAGAGTTTATPATG	731
Db	706	CTTTATTTTGCATTTTGGTTTCAAGCATATGATCTGAAGAAGTGGGACCAAGTTTATAGT	765
Qy	732	AAGGCATTTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAAGCTGAGATATGGCTA	791
Db	766	GAAGTTCTTAGGAAGACCCACTACATTTATTTGAGACAATGGGAAGCTGAATGTGGCTC	825
Qy	792	ATACGAACATATTGGGATTTTGAATTTCCCTCAACCATTACCAACCTAACTTTAGTTTGT	851
Db	826	ATTCGAACCTATTGGGATTTTGAATTTTCCCTCGGCCATTCTTACCAATGTGTATTTGTT	885
Qy	852	GGAGGATTCGACTGTAAACCTGCCAAAGCTTTGCCTAAGGAAATGGAAAATTTTGTCCAG	911
Db	886	GGAGGACTTCACCTGTAAACCAGCCAAACCCCTTGCTTAAGGAAATGGAAGAGTTTGTGCG	945
Qy	912	AGTTTCAGGGGAAGATGATTTGTGGTGTTTTCTCTGGGGTCACTGTTTCAAAATGTTACA	971
Db	946	AGCTCTGGAGAAAATGGTATTGTGGTGTTTTCTCTGGGGTGCATGATGATCAACATGCTCA	1005
Qy	972	GAAGAAAGGCTAATATTCATTTGCTTACGCCCTTTGCCAGATCCACAGAAGGTGTTATGG	1031
Db	1006	GAAGAAAGTCCAAACATGATTTGCATCAGCCCTTGCCAGATCCCAAAAAGGTTCTATGG	1065
Qy	1032	AGGTACAAAGGAAAAAACCATCCACATTAGGAGCCCAATACTCGGTGTATGATGGATA	1091
Db	1066	AGATTTGATGGCAAGAGCCAAATACTTTAGTTTCCAATACTCGACTGTATAAGTGGTTA	1125
Qy	1092	CCCCAGAAATCATCTTTGTGTATCCCAAAACCAAAAGCTTTTATCACTCATGTGGGAATG	1151
Db	1126	CCCCAGAAATGACCTTCTTGCTATCCCAAAACCAAAAGCTTTTAACTCATGTGGGAACC	1185
Qy	1152	AATGGGATCATGAAGCTATTTTACCATGGGTCCCTATGTGGGAGTTTCCCATATTTGGT	1211
Db	1186	ANTGGCATCATGAGGCGATCTACCATGGGATCCCTATGTGGGCACTCCCTGTGTTTGGC	1245
Qy	1212	GATCAGCTTGATAACATAGCTCATATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTC	1271
Db	1246	GATCAACATGATAACATTTGCTCATATGAAAGCCAAAGGAGCAGCCCTCAGTGTGGACATC	1305
Qy	1272	AAAACATAGACAAGCGAAGATTTTACTGAGGGCTTTTGAGAACAGTCAATTACCGATTCCTCT	1331
Db	1306	AGGACCATGTCAAGTAGAGATTTGCTCAATGTCATTTGAAGTCAGTCAATTAAGCCCTATC	1365
Qy	1332	TATAAAGAGAAATGCTATGAGATTAATCAAGAAATTCACCATCATCAACCTGTAAAGCCCTA	1391
Db	1366	TATAAAGAGAAATATCATGAAATTTATCAAGAAATTCATCATGATCAACCGGTGAAGCCCTG	1425
Qy	1392	GATCGAGCACTTCTGTGATTCGATTTTGTTCATGCGCCACAAAGAGGACCAAGCACTCGGA	1451
Db	1426	GATCGAGCACTTCTGTGATTTGATTTGTCATGGCCATTAAGGACCAAGCACTTCGG	1485
Qy	1452	TCAGCTGCCCATGACCTTCACTGGTTTCCACGACTACTCTATAGATGTGATTTGGGTTCCTG	1511

Db	1486	GTCCGAGCCCAACCTCAGCTGGATCCACTCTTTGGATGTAGATATTCCTG	1545
Qy	1512	CTGACCTGTGTGGCACTGCTATATCTTCTTTCACAAAATGTTTTTATTTTTCTCTGTCAA	1571
Db	1546	CTGCGCTGCGTGCAACTATGATATTTATCATCAAAAATGTTTGCCTGTTTTTGTGTTCCGA	1605
Qy	1572	AAATTTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTTCCAAA	1621
Db	1506	AAGCTTCCCAAAACAGGAAAGAGAGAAAAGGATTTAGTTATATCAAAA	1655

RESULT 13	
HSU59209	
LOCUS	2107 bp mRNA linear PRI 02-JUL-1998
DEFINITION	Homo sapiens C19steroid specific UDP-glucuronosyltransferase mRNA, complete cds.
ACCESSION	U59209
VERSION	U59209.1 GI:3287472
KEYWORDS	UCT2B17G.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
AUTHORS	1 (bases 1 to 2107)
TITLE	Beaulieu,M., Levesque,E., Hum,D.W. and Belanger,A. Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids
JOURNAL	J. Biol. Chem. 271 (37), 22855-22862 (1996)
MEDLINE	96394358
PUBMED	8798464
REFERENCE	2 (bases 1 to 2107)
AUTHORS	Hum,D.W., Beaulieu,A., Beaulieu,M. and Levesque,E.
TITLE	Direct Submission
JOURNAL	Submitted (23-MAY-1996) Laboratory of Molecular Endocrinology,

BASE COUNT	657 a	382 c	424 g	644 t
ORIGIN	ITRCCLLFCFRKKLARTGKKKKKK			
Query Match	47.4%; Score 775.6; DB 9; Length 2107;			
Best Local Similarity	68.6%; Pred. No. 6.8e-190;			
Matches 1104; Conservative	0; Mismatches 494; Indels 12; Gaps 2;			

CDS

QY	141	GTCAAGGTCATCTAGAAAGCTCATAGTGAGAGGCCATGAGTAACAGTATTGACTCATC	200
Db	166	ATGAAGACAATCTCGAAGAGCTTGTTCAGAGGGTCATGAGTGATTTGTTGACATCT	225
QY	201	TCAAGCCCTTCGTTAATTGACTACAGGAAGCCTTCTGCATTTGAATTTGAGTGTTCCAT	260
Db	226	TCGGCTTCATTCCTGCAATGCCAGTAATCATCTGCTATTAAATTAAGAAGTTATCCT	285
QY	261	ATGCCACAGGACAGACAGAGAAATGAAATATTT-----GTTGACCTAGTCTCG	311
Db	286	ACATCTTTAACTAAAAATGATTTGGAAGATTTTTTATGAAATGTTTCATAGATGGACA	345
QY	312	AATGTCTTCCGAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATGATTTTTTGT	371
Db	346	TATAGTATTTCAAAAATACATTTTGGTCATATTTTTCACAACTACAAGAAATGTTGG	405
QY	372	GAAATAGAGGAACTTTAAAAATGATGTGTGAGAGCTTTATCTACAATCAGACGCTTATG	431
Db	406	GAATATCTTGACTATAATATAAGCTCTGTGAAGATGCGAGTTTGNACAGAAACTATG	465
QY	432	AAGAAGCTACAGAAACCAACTACGATGTAATGCTTTATAGACCCCTGTGATCCCTGTGGA	491
Db	466	AGAAACTACAAGGTCAAAATTTGATGCTCTTGGCAGATGCGCGTTAAATCCCTGTGCT	525
QY	492	GACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTTAGAATTTCTCTAGA	551
Db	526	GAGCTGCTGGCTGAACACTTAACATACCCTTCTGTACAGTCCCGCTTCTCTGTTGGC	585
QY	552	GCAATATGGAGGAGCTGTGGGAAACTTCAGCTCCACTTTCCATGTACCTGTGCGCT	611
Db	586	TACACAGTTGAGAAAGTGGTGAGGATTTCTGTCCCTCTCTCTATGTACCTTGTCTT	645
QY	612	ATGACAGGACTAACAGACAGATGACCTTCTCGAAAGAGTAAAAAATCAATGCTTCA	671
Db	646	ATGTCAGAATTAAGTGATCAAAATGATTTTCATGGAGAGATAAAAATATGATATATG	705
QY	672	GTTTTGTCCACTCTCGATTCAGGATTCAGACTATCAATTTTGGGAAGAGTTTTATAGT	731
Db	706	CTTTATTTTGACTTTTGGTTTCAAGCATATGATCTGAAGAAAGTGGGACCACTTTTATAGT	765
QY	732	AAGCATTTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGATATGCTA	791
Db	766	GAAGTTCTAGCAAGACCACACTACATTTATGAGACATGGGAAGAGTGAATGTGGCTC	825
QY	792	ATACGAACATATTTGGGATTTTGAATTTCCCTCAACCATACCAACTAATTTGAGTTTGT	851
Db	826	ATTGGAACCTATTTGGGATTTTGAATTTTCCCTGCGCCATCTTACCAAAATGTTGTTGT	885
QY	852	GGAGATTGCACTGTAACCTGCCAAAGCTTTGGCTAAGGAATGGAATTTTGTGTCAG	911
Db	886	GGAGACTTCACTGTAACCCAGCCAAACCTTGCCTAAGGAATGGAAGAGTTTGTGCGAG	945
QY	912	AGTTCAGGGAGATGTTATTTGGTGTCTTCTGCGGTCACTGTTTCAAAATGTTTACA	971
Db	946	AGCTCTGGAGAAATGTTATTTGGTGTCTTCTGCGGTGCGATGATCACTAATGCTCA	1005
QY	972	GAAGAAAGGCTAATATCATTTGCTTACGCCCTTGGCCAGATCCCAAGAGGTGTTATGG	1031
Db	1006	GAAGAAAGTCCACATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG	1065
QY	1032	AGGTACAAAGGAAAAAACCATCCACATAGGAGCCAAATCTCGGTGATGATGATGATA	1091
Db	1066	AGATTTGATGGCAAGAGCCAAATACTTTAGTTTCCAATCTCGACTGATATAGTGGTTA	1125
QY	1092	CCCAGAAATGATCTTCTTGTGTCATCCCAAAACCAAGCTTTTATCACTCATGTTGGAATG	1151
Db	1126	CCCAGAAATGATCTTCTTGTGTCATCCCAAAACCAAGCTTTTATCACTCATGTTGGAATG	1185
QY	1152	AATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGTTGGGAGTCCCATATTTGGT	1211
Db	1186	AATGGGATCTATGAGGCGATCTACCATGGGATCCCTATGTTGGGATCCCTTGTGTTGG	1245
QY	1212	GATCAGCTTGTATACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATPAAACTTC	1271

[illegible]

```

RESULT 14
AF112112
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AF112112
Macaca fascicularis UDP-glucuronosyltransferase 2B19 precursor,
mRNA, complete cds.
AF112112.1 GI:4580601
Macaca fascicularis.
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 2108)
Belanger, G., Barbier, O., Hum, D.W. and Belanger, A.
Molecular cloning, expression and characterization of a monkey
steroid UDP-glucuronosyltransferase, UGT2B19, that conjugates
testosterone
Eur. J. Biochem. 260 (3), 701-708 (1999)
99203465
10102998
2 (bases 1 to 2108)
Belanger, G., Barbier, O., Hum, D.W. and Belanger, A.
Direct Submission
Submitted (07-DEC-1998) Molecular Endocrinology, CHUL Research
Center, 2705 Laurier Blvd, Sainte Foy, Quebec G1V 4G2, Canada
Location/Qualifiers
1..2108
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
26..1612
/EC_number="2.4.1.17"
/note="UGT2B19"
/codon_start=1
/product="UDP-glucuronosyltransferase 2B19 precursor"
/protein_id="AAD24435.1"
/db_xref="GI:4580602"
/translation="MSMKWTSALLLIQLCYLSFGSGKVLVMPTEFSHWNKTKILD
ELVGRHEVTVLAYSTLTPDPNPPSLKFEICPTSLTETEFQDSVTLVKRWSDIRK
DTWPFHFLVQEMWTGDMTRKCKDQVSNKKLWKQLQESRDFVLADAI SPGELL
AELLKIPVLSRPGVALEKHGGGFLFPSPYVPTMSLRQMTFMERVQNKIYMY
YDFWFOVDVKNRDSQYKVLGRPTTLFPEIMAKAEILRNRTWDFQFPHLPNVEF
VGLGHCPAKPLPKMEBEFVQSGDGVVFSLGSVMVSNSEERANVIASALAKIPQK
VWLPFDGNKTLNGLQYKWLQPNDDLGHKPTAFITHGGANGIYEATIHGIPMVG
VLPFADQPNTHAHKGAARVLDLNLNLTVDINDPIYKENAMKLSLSIHH"

```


Qy 55 TTCTGCTCCTGACGCTCTTCTGCTGTGGCTGTGGATCTCTGCGAAGATCCCTGGTGTGGC 114
Db 8 TGCTGATACAGCTCAGCTGCTTACTTGTAGCTCTGGAAGCTGTGGAAGGCTGCTAGTGTGGC 67
Qy 115 CTTGTGACATGAGCCATTTGCTTAATGTCTCAAGCTCATTTCTAGAAGAGCTCATAGTGCAGAG 174
Db 68 CCACAGAATACAGCCATTTGATAAATGAAGACAATCTCTGGAAGAGCTTGTTCAGAGG 127
Qy 175 GCCATGAGGTAACAGTATTGACCTCAAGGCTTCTGTTAAATGACTACAGAGAGCCCT 234
Db 128 GTCATGAGGTGACTGTGTGACATCTTCGGCTTCTACTCTTGTCAATGCGACGTAATCAT 187
Qy 235 CTGCATTGAAATTTGAGGTGGTCCATATGCC-----ACAGGACAGCAAGAGAA 285
Db 188 CTGCTATTAATAGAAAGTTTATCTTACATACTTTTAACTAAATTTATTGGGAAGATCTTC 247
Qy 286 ATGCAATATTTGTTGACCTAGCTCGAATGTCTTGGCAGGCTTATCAACCTGGCAATCAG 345
Db 248 TTCTGAAATCTCGATAGATGGATATATGTTTCAAAAAATACATTTTGGTCATATT 307
Qy 346 TTATAAATTAATGATTTTTTTGTTGAAATAAGAGGAACTTTAAAAATGATGTGTGAGA 405
Db 308 TTTCACAATTAACAAGAAATTTGTTGGGCATATTGACTACAGTACAGGCTCTGTAAG 367
Qy 406 GCTTTATCTACAACTCAGAGGCTTATGAAGAGCTACAGGAACCAACACTACGATGTAATGC 465
Db 368 ATGCAAGTTTGAATAAGAAACTTATGATGAACACTACAAGAGTCAAGATTTGATGTCAATC 427
Qy 466 TTATAGACCTGTGATTCCTCTGGAGACCTGATGGCTGAGTGTCTTCCAGTCCCTTTTG 525
Db 428 CTGGAGATCCCTGTTTCCCTGTGTGAGCTACTGTACAGGCTACTTAAACATAACCCCTTG 487
Qy 526 TGCTCACACTTGAATTTCTGTGAGGCAATATGAGCGAAGCTGTGGGAACCTTCCAG 585
Db 488 TCTACAGTCTCGCATCTCTGTGGCTACACAATTGAGAAGATGGTGGAGGATTTCTGT 547
Qy 586 CTCCACTTTCTATGCTGCTGTGCTATGACAGGACTTAACAGACAGATGACCTTCTGG 645
Db 548 TCCCTCTCTGATGCTGCTGTTATGTCAGAAATTAATGATCAAAATGATTTTCTATGG 607
Qy 646 AAAGATTAATAATTCATGCTTTTCAGTTTGTCTCCACTTCTGGATTCAGGATACGACT 705
Db 608 AGAGATTAATAATATGATACATATGCTTTATTTTGTGACTTTTGGTTTCAAAATTTATGTC 667
Qy 706 ATCATTTTGGGAAGACTTTTATAGTAAGGATTTAGGAAGCCCACTACATTTATGTGAGA 765
Db 668 TGAAGAAGTGGACAGCTTTTATAGTGAAGTCTTAGGAAGACCCACTACATTTATGAGA 727
Qy 766 CTGTGGGAAAGCTCAGATATGGCTAATACGAACATATGGGATTTTGAATTTCTCTCAAC 825
Db 728 CAATGGGAAAGCTGAATGTGGCTCATTCGAACCTATGGGATTTTGAATTTCTCTGGCC 787
Qy 826 CATACCAACCTTACTTTGAGTTTGTGGAGATTTGCACTGTAAACCTGCCAAAGCTTTGC 885
Db 788 CATTTCTACCAAAATGTTGATTTTGTGGAGGACTTCACTGTAAACAGCCAAACCCCTGC 847
Qy 886 CTAAGAAATGGAATAATTTGTCCAGAGTTTCAGGGAAGATGTTATTTGGTGTGTTTCTC 945
Db 848 CTAAGGAATTTGGAAGAGTTTGTGCAGAGCTCTGGAGAAATGTTATGTTGGTGTGTTTCTC 907
Qy 946 TGGGTCACCTGTTTCAAAATGTTACAGAGAAAGGCTAATATCATTTGCTTCCAGCCCTTG 1005
Db 908 TGGGTCGATGATCAGTAACATGTACAGAGAAAGTGCACACATGATTGTCATGAGCCCTTG 967
Qy 1006 CCGAGATCCACAGAGGCTTTATGAGGTACAAAGGAATAAACCATCCACATTAGGAG 1065
Db 968 CCGAGATCCCAAAAGGTTCTATGAGATTTGATGGCAAGAGCCAAATACTTTAGGTT 1027
Qy 1066 CCAATCTCGGCTCTATGATTTGATACCCAGATGATCTTCTTGGTCATCCCAAAACCA 1125
Db 1028 CCAATCTCGACTGTACAAGTGTATACCCAGATGACCTTCTTGTCTATCCCAAAACCA 1087
Qy 1126 AAGCTTTTATCCTCATGTTGGGAATGAATGGGATCTATGAAGCTATTTTACCATGGGGTCC 1185

Db 1088 AAGCTTTTATTAATCTCATGGTGAACCAATGGCATCTATGAGGCGATCTACCATGGATCC 1147
Qy 1186 CTATGGTGGAGTTCCCATATTTTGTGTGATCAGCTTTGATAAACAATAGCTCACATGAAGGCCA 1245
Db 1148 CTATGGTGGGCAATCCCTTGTGTGGGATCAAGATGATAACATTTCTCACATGAAGGCCA 1207
Qy 1246 AAGGAGCAGCTGTAGAAAATAAACTTCAAAAGCTATGACAAGCGAAGATTTTACTGAGGGCTT 1305
Db 1208 AGGAGAGACCCCTCAGTGTGACATCAGACCATGTCAAGTAGAGATTTGCTCAATGCAT 1267
Qy 1306 TGAGAACAGTCATTAACCGATTCTCTTTATAAAGAGAAATGCTATGAGATTATCAAGAAATC 1365
Db 1268 TGAAGTCAGTCATTAATGACCCAGCTCTATAAAGAGAAATGCTATGAATTTATCAAGAAATC 1327
Qy 1366 ACCATGATCAACCTGTAAAGCCCCCTCTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCTATGC 1425
Db 1328 ATCATGACCAACCAATGAAGCCCCCTGGATCGAGCAGTCTTCTGGATTCAGTTTGTCTATGC 1387
Qy 1426 GCCACAAAGSAGCCAGCACCTGCGATCAGCTGCCATGACCTCACCTGCTGTTCCAGCACT 1485
Db 1388 GCCACAAAGSAGCCAGCACCTTTCGAGTCCGAGCTCACACCTCACCTGGATCCAGTACC 1447
Qy 1486 ACTCTATAGATGATTTGGGTTCTCTGCTGACCTGTGTGGCAACTGTCTATTTCTTTGTTCA 1545
Db 1448 ACTCTTTGGATGTGATAGCATTTCTGCTGGCCTGGTGGCAACTGTCTATTTATCATCA 1507
Qy 1546 CAAAATGTTTTTTTATTTTCTGTCAAAAATTTAATAAACTAGAAAGATAGAAAGAGGG 1605
Db 1508 CAAAATTTTGGCTGTTTTTTCGAAAGCTTGCCAAAAAAGGAAAGAGAAAGAG 1567
Qy 1606 AATAGATCTTTCCAAA 1621
Db 1568 ATTAGTTTATATCAAAA 1583

Search completed: July 16, 2003, 13:28:03
Job time : 2983.96 secs

OY 1274 AACTATGACAGCGAAGATTACTGAGGGCTTTGAGAACAGTCATTACCGATTCCCTCTTA 1333
 DB |||||||
 OY 1288 AACTATGACAGCGAAGATTACTGAGGGCTTTGAGAACAGTCATTACCGATTCCCTCTTA 1347
 DB |||||||
 OY 1334 TAAAGAGATGCTATGAGATTATCAAGAATTCACCATGATCAACCTGTAAGCCCTAG 1393
 DB |||||||
 OY 1348 TAAAGAGATGCTATGAGATTATCAAGAATTCACCATGATCAACCTGTAAGCCCTAG 1407
 DB |||||||
 OY 1394 TCAGCAGTCTCTGATCGAGTTTCTCATGCCCAAGAGGCAACCTCGGATC 1453
 DB |||||||
 OY 1408 TCAGCAGTCTCTGATCGAGTTTCTCATGCCCAAGAGGCAACCTCGGATC 1467
 DB |||||||
 OY 1454 AGTGCCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGATGGTCTCTGCT 1513
 DB |||||||
 OY 1468 AGTGCCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGATGGTCTCTGCT 1527
 DB |||||||
 OY 1514 GACCTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTATTTCTGCAAAA 1573
 DB |||||||
 OY 1528 GACCTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTATTTCTGCAAAA 1587
 DB |||||||
 OY 1574 ATTTAATAAACTAGAAAGATAGAAAGAGGGAATAGATCTTTCCAAATTCAGAAAGAC 1633
 DB |||||||
 OY 1588 ATTTAATAAACTAGAAAGATAGAAAGAGGGAATAGATCTTTCCAAATTCAGAAAGAC 1647
 DB ||||
 OY 1634 CTG 1636
 DB ||||
 OY 1648 CTG 1650

RESULT 5

ID AAS91380
 AC AAS91380
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #27184.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.*
 DR P-PSDB; ABG27193.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID No 27184; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful in treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1766 BP; 542 A; 359 C; 363 G; 502 T; 0 other;

Query Match 50.2%; Score 821.2; DB 23; Length 1766;
 Best Local Similarity 71.9%; Pred. No. 7.4e-225;
 Matches 1107; Conservative 0; Mismatches 418; Indels 15; Gaps 2;
 OY 94 GTGGGAAGCTCCTGGTGGCCCTGTGACATGAGCCATGGCTTAATGTCAAGGTCAATC 153
 DB |||||||
 OY 122 GTGGGAATGTTTGAATTTGGCCCAATGGAAGGTAGTCATTGGCTAAATGTTAAGATAATTA 181
 DB |||||||
 OY 154 TAGAAGAGCTCATAGTCAGAGGCCATGAGGTACAGTATTGACTCACTCAAGGCCCTCGT 213
 DB |||||||
 OY 182 TAGATGAGTCATTAAAAGGAGCATTAATGTGACTGTCTCTGCTCTGGTGCACATTT 241
 DB |||||||
 OY 214 TAATTGACTACAGGAAGCCCTTCTGCAATTTGAGGTGGTCCATATATACAGGTCCTTGGCA 273
 DB |||||
 OY 242 TCATCACACCAACCTCTAACCCATCTCTGACATTTGAAATATATAAGTGCCTTGGCA 301
 DB |||||
 OY 274 GAACAGAAAGAAATGAAATATTTTGGACCTCTGCTGAATGCTTTCGCCAGGCTTATCAA 333
 DB |||||||
 OY 302 AAGAAAGATAGAGGAGTAATTAAGGACCTTCGTTTCGACATGGCTGGAAATATAGACCAT 361
 DB |||||||
 OY 334 CTTGGCAATCAGTTATAAATTAATTAATTTTGTGTTGAAATAGAGGAACCTTTAAAAA 393
 DB |||||||
 OY 362 CTCCTTCAACCAATTTGGAGATTTCTATCAGGAGATGCGCAAGTAATCAAGGACTTCCACA 421
 DB |||||||
 OY 394 TGATG-----TGTGAGAGCTTTATCTACAAATCAGAGCTTATGAGAACTAC 441
 DB |||||||
 OY 422 TGGTGTCTCAGGAGATCTGTGATGGCGTCTTAAAAACCAACAGCTGATGGCAAGCTAA 481
 DB |||||||
 OY 442 AGGAAACCAACTACGATTAATGCTTTATAGACCTCTGATTCCTCTGGAGACCTGATGG 501
 DB |||||||
 OY 482 AGAAAGCAAGTTTGAAGTCTGCTGATCCAGTATTTCTTGTGGCGATATAGTAG 541
 DB |||||||
 OY 502 CTGAGTTGCTTCAGTCCCTTTTGTGCTCAGCATTTAGAATTTCTGTAGAGGCAATATGG 561
 DB |||||||
 OY 542 CTTTAAACTTGGAAATTCATTTATGTTACTCTCTGAGGTTTCTCCAGCCTCAACAGTGG 601
 DB |||||||
 OY 562 AGCAGAGCTGTGGGAAACTTCCAGCTCCACTTTCTATGTATCTGCTGCTATGACAGNAC 621
 DB |||||||
 OY 602 AAAAGCACTGTGGGAGGTACCATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 661
 DB |||||||
 OY 622 TAACAGACAGAAATGACCTTTCTGGAAGAGTAAAAAATTTCAATGCTTTTGTGTTTCC 681
 DB |||||||
 OY 662 TCACCGACCAATGTCTTCTCACTGACAGATAAGAAATTTTCATCTCTCTACCACTAC 718
 DB |||||||
 OY 682 ACTTCTGGATTCAGGATTCAGCACTATCATTTTGGGAAGAGTTTTATAGTAAGCAATAG 741
 DB |||||||
 OY 719 AGGACTACATGTTTGAAGAACTCTTTGGAATCATGGATTCTACTATATAGTAAGCTTAG 778
 DB |||||||
 OY 742 GAAGCCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGATATGCTATGCTATGACACAT 801
 DB |||||||
 OY 779 GAAGACCCACTACGTTATGTGAGACTATGGGAAAGCTGAAATTTGGTTAATCGGAACAT 838
 DB |||||||
 OY 802 ATTGGGATTTGAATTTCTCTCAACCATACCAACCTTAACCTTTGAGTTTGTGGAGGATTGC 861
 DB |||||||

839	ATTGGGATTTTGAATTTCTCCTGCCATACATTACCTAAATTTTGGAGTATGC	898
862	ACTGTAACACTGCCAAAGCTTTGCCTAAGGAAATGGAAAAATTTTGTCCAGAGTTCAGGGG	921
899	ACTGCAAACTGCCAAACCTTTACCTAAGGAAATGGAAGAAATTTATCCAGAGCTCAGGTA	958
922	AAGATGGTATTGGGTGTTTTCTCTGGGGTCACCTGTTTCAAATGTTTACAGAGAAAAAGG	981
959	AAAAATGGTGTGGGTGTTTTCTCTGGGGATCAATGTTCAAAAACCTTTACAGAGAAAAAGG	1018
982	CTAATATCATTTGCTTTCAGCCCTTGCCCCAGATCCACAGAAAGGCTTATGGAGTACAAAG	1041
1019	CCAATCTTATGGCTCAGCCCTTGCCCCAGATCCACAGAGGTTTTTATGGAGATACAAAG	1078
1042	GAAAAAAACCATCCACATTAAGGACCAATACTCGGCTGTATGATTGGATACCCCAAGATG	1101
1079	GAAGAAACAGCCACATTAAGGAACAATACTCAGCTCTTTGATTGGATACCCCAAGATG	1138
1102	ATCTTCTTGTGTATCCCAAAACCAAGCTTTTATCACTCATGTGTGGAATGAATGGGATCT	1161
1139	ATCTTCTTGACATCCCAAAACCAAGCTTTTATCACTCATGTGTGGAATGAATGGGATCT	1198
1162	ATGAAGCTATTTACCATGGGGTCCCTATGTGGGAGTTCACATATTTGGTGATCAGCTTG	1221
1199	ACGAAGCTATTTACACGGAGTCCCTATGTGGGAGTTCACATGTTTGTCTGATCAGCCTG	1258
1222	ATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAAACTATGA	1281
1259	ATAACATTGCTCACATGAAGGCCAAAGGAGCAGCTGTGGAAGTGAACCTTAAACACAATGA	1318
1282	CAAGCGAAGATTTACTGAGGGCTTTGAGACAGTCATTACCGATTCTCTTATTAAGAGAGA	1341
1319	CAAGTGTGGATTTGCTTTAGCGCTTTGAGACAGTCATTATGAACCTTCTTATTAAGAGAGA	1378
1342	ATGCTATGACATTTACAAGAAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGCAG	1401
1379	ATGCTATGAGTTATCAAGAATTTACCCATGATCAACCTGTAAAGCCCTGGATCGACGAG	1438
1402	TCTTCTGGATCGAGTTTGTCTATGGCCACAAAGGAGCCAAAGCCTCGCATCAGCTGCC	1461
1439	TCTTCTGGATCGAGTTTGTCTATGGCCACAAAGGAGCCAAAGCCTTCGGGTTTCAGGCC	1498
1462	ATGACCTCAGCTGGTTCAGCAGTACTCTATAGATGTGATTGGGTTCTCTGCTGACCTGTG	1521
1499	ATGACCTCAGCTGGTTCAGTACACATCTTTGGATGTAATGGGTTCTCTGCTGCTGTG	1558
1522	TGGCAACTGCTATATTCTGTTCAAAAATGTTTTTATTTTCTCTGTCAAAAAATTAATA	1581
1559	TGACAAGGCTATATTTTGGTCATACAAATGTTGTTTCTCTGTCAAAAATTTGGTA	1618
1582	AAACTAGAAAGATAGAAAGAGGGAATAGATCTTTCCAAA	1621
1619	AGATAGAAAAGAGAAAAAAGAGAAATAGGTCAAGAAAAA	1658

DEC 11 1964

RESULT 6
ABK84210

ABK84210
ID ABK84210 standard: cDNA: 2090 bp.

XX
OTRACOMI

AC ABK84210;

XX
DT 14-AUG-2002 (first entry)

XX	Human cDNA differentially expressed in granulocytic cells #781.
XX	
DE	Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX	viral infection; parasitic infection; protozoal infection;
KW	fungal infection; sterile inflammation; disease; psoriasis;
KW	reumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW	adult respiratory distress syndrome; inflammatory bowel disease
KW	Crohn's disease; ulcerative colitis; periodontal disease;
KW	granulocyte activation; chronic inflammation; allergy.

XX	Homo sapiens.
OS	
XX	WO200228999-A2.
PN	
XX	11-APR-2002.
PD	
XX	03-OCT-2001; 2001WO-US30821.
PF	
XX	03-OCT-2000; 2000US-237189P.
PR	
XX	(GENE-) GENE LOGIC INC.
PA	
XX	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
PI	
XX	WPI; 2002-435328/46.
XX	
DR	
XX	Detecting granulocyte activation by detecting differential expression
PT	of genes associated with granulocyte activation, which serves as
PT	diagnostic markers that is useful for monitoring disease states and
PT	drug toxicity -
XX	
PS	Claim 1: SEQ ID No 781: 114pp: English.

Query Match	47.8%	Score 782;	DB 24;	Length 2090;
Best Local Similarity	68.8%;	Pred. No. 1.4e-213;		
Matches 1108: Conservative	0;	Mismatches 490;	Indels 12;	Gaps 2;

Qy	24	ATCATGAGTCTGACAAAGTCAGCTTTGGTAATTTCTGCTCTGACGCTCTCTCTGT	--GTT	80
Db	16	ACCAGGATGCTCTGAAATGGACGTCAGTCTTCTGCTGATACAGCTCAGTGTGTTACTTT		75
Qy	81	GGCTGTGGATTCTGTGGAAAGTCCTGGTGTGCCCTGTGACATGACCCATTTGGCTTAAT		140

Db	76	AGCTCTGGAAGCTGTGAAAGGTGCTAGTGTGGCCACAGAAATACAGCCATTGGATAAAT	135
QY	141	GTCAAGGTCATTCTAGAAAGCTCATAGTGAGAGGCCATGAGTAACTATTTAGTCTAC	200
Db	136	ATGAAGACAATCTTGAAGAGCTTGTTCAGAGGGTCATGAGTGACTGTGTGACATCT	195
QY	201	TCAAAGCCTTCGTTAATTGACTACAGGAAGCCCTTCTGCATTTGAAATTTTTCAGGTGGTCCAT	260
Db	196	TGGGCTTCTACTCTTCTCAATGCCAGTAATCATCTGCTATTAAATTAAGAAGTTATCTCT	255
QY	261	ATGCC-----ACAGACAGACAGAGAAATGAATATTTGTTGACCTAGTCTG	311
Db	256	ACATCTTTAACTAAATAATGTTGGAGATTTCTTTCTGAAATTCCTCATAGATGGATA	315
QY	312	AATGTCTTCCAGGCTTATCAACCTGGCAATCAGTTTATAAATTAATGATTTTGT	371
Db	316	TATGGTGTTCAAAAATACATTTTGGTCATATTTTTCACATTAACAAGATTTGTTGG	375
QY	372	GAAATGAAGAGAACTTTAAATAATGATGTGTGAGAGCTTTTATCTACAATCAGACGCTTATG	431
Db	376	GAAATATTAGTACTACAGTAACAAGCTCTCTAAAGATGCGAGTTTGAATAAAGAACTTATG	435
QY	432	AAGAAGCTACAGAAACCACTACGATTAATGCTTATAGACCTGTGATTCCTCTGGA	491
Db	436	ATGAAGACTACAGAGTCAAGTTTGAITGTCATCTTGGCAGATGCCCTTAACTCCCTGTGT	495
QY	492	GACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAATTTCTGTAGGA	551
Db	496	GAGTACTGGCTGAACATTTTAACATACCTTTCTGTACAGCTTCGATTCCTGTGGC	555
QY	552	GGCAATATGGAGGGAAGCTTGGGAACTTCCAGCTCCACTTTCCCTATCTACTGTGCT	611
Db	556	TACACATTTGAGAAGAATGCTGGAGGATTTCTGTTCCCTCTCCCTATGTACCTGTGTT	615
QY	612	ATCAGAGGACTACACACAGATGACCTTTCTGGAAAGAGTAAAAAATCAATGCTTTCA	671
Db	616	ATGTCAGAATTAAGTCAATCAATGATTTTCATGGAGAGGATAAAAAATATGATACATATG	675
QY	672	GTTTTCTCCACTTCTGGGATTCAGGATTAAGTATGCTATTTTGGGAAGTTTTATAGT	731
Db	676	CTTTATTTTGACTTTTGGTTTCAATTTATGATCTGAAGAGTGGGACCATGTTTATAGT	735
QY	732	AGGCATATAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGATAGGCTA	791
Db	736	GAAGTTCTAGGAAGACCCACTACATTTATTGAGACAAATGGGAAAGCTGAAATGTGGCTC	795
QY	792	ATACGAACATATTTGGGATTTGAAATTTCTCAACCATACCACTTAACCTTTGATTTGTT	851
Db	796	ATTCGAACCTATTTGGGATTTTGAATTTCTCGGCCATTTCTTACCAATGTTGATTTGTT	855
QY	852	GGAGGATTCACCTGTAAACCTGCAAGCTTTTCCCTAAGGAATGGAATTTTGTCCAG	911
Db	856	GGAGGACTTCACTGTAAACCAAGCCCAACCCCTGCTTGAAGAAATGGAAGATTTTGCAG	915
QY	912	AGTTCAGGGAAGATGGTATTTGGTGTGTTTCTCTGGGGTCACTGTTTCAAAATGTTACA	971
Db	916	AGCTCTGGAGAAATGGTATTTGGTGTGTTTCTCTGGGGTCACTGATAGTAACTATGCA	975
QY	972	GAAGAAAGGCTAATATCATCTGCTCAGCCCTTGGCCAGATCCACAGAGGTGTTATGG	1031
Db	976	GAAGAAGTGCCAACTATGATGATCAGCCCTTGGCCAGATCCCAAGAGTGTCTATGG	1035
QY	1032	AGGTACAAAGGAAAAAACCATCCACATTAGAGCCCAATACTCGGCTGTATGTTGGATA	1091
Db	1036	AGATTTGATGGCAAGCAACCAATACATAGGTTCACATCTCGACTGTACAAGTGTTA	1095
QY	1092	CCCCAGATGATCTTCTTGGTCAATCCCAAAACCAAGCTTTTATCACTCATGTGGAATG	1151
Db	1096	CCCCAGATGACCTTCTTGGTCAATCCCAAAACCAAGCTTTTATCACTCATGTGGAAC	1155
QY	1152	AATGGGATCATGAAGCTATTTACCATGGGGTCCCTATGTTGGGAGTTTCCCATATTTGGT	1211
Db	1156	AATGGCATCTATGAGGCGATCTACCATGGGATCCCTATGTTGGGCAATCCCTGTTTGG	1215
QY	1212	GATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAATTC	1271
Db	1216	GATCAACATGATAACATGCTCATGTAAGCCAAAGGAGCAGCCCTCAGTGTGGACATC	1275
QY	1272	AAACTATGACAAAGGAAGATTTTACTGAGGGCTTTTGAACACAGTCAATACCGATTCCTCT	1331
Db	1276	AGGACCATGCTCAAGTAGAGATTTTGTCTCAATGCAATGAAGTCACTAATGACCTGTC	1335
QY	1332	TATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTCTAAAGCCCTA	1391
Db	1336	TATAAGAGAAATGCTATGAAATTTATCAAGAAATTCATGACCAACCAATGAAGCCCTG	1395
QY	1392	GATCGAGCAGTCTTCTGGATCGAGTTTGTATCGCCACAAAGGAGCCCAAGCCTGCGCA	1451
Db	1396	GATCGAGCAGTCTTCTGGATCGAGTTTGTATCGCCACAAAGGAGCCCAAGCCTGCGCA	1455
QY	1452	TCAGCTGCCATGACCTCACCTGGTTCAGCACTACTCTATAGATGTATGGTTCCTG	1511
Db	1456	GTGCGAGCTCACAAACCTCACCTGGATCCAGTACCACCTTTTGGATGTGATGATTCCTG	1515
QY	1512	CTGACCTGTGTGCAACTGCTATATTTTGTTCACAAATGTTTATTTTCTCTGCA	1571
Db	1516	CTGGCTGCTGTGCAACTGCTATATTTATCATCACAAATTTTGCCTGTTTGTTCCTGCA	1575
QY	1572	AAATTTTAAATAAGTAGAAGAGATAGAAAAGAGGGAATAGATCTTTTCCAAA	1621
Db	1576	ANGCTGCCAAACAGGAAGAAAGAGAAAGAGATTTATATATATCAAAA	1625
RESULT 7			
ID	ABN96795	ABN96795 standard; DNA; 2090 BP.	
XX	AC	ABN96795;	
XX	DT	13-AUG-2002 (first entry)	
XX	Gene	#3293 used to diagnose liver cancer.	
XX	Gene:	liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KW	metastatic	liver tumour; cytostatic; expression profile; disease state;	
KW	disease	progression; drug toxicity; drug efficacy; drug metabolism.	
XX	OS	Homo sapiens.	
XX	PN	WO200229103-A2.	
XX	PD	11-APR-2002.	
XX	PF	02-OCT-2001; 2001WO-US30589.	
XX	PR	02-OCT-2000; 2000US-237054P.	
XX	PA	(GENE-) GENE LOGIC INC.	
XX	PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX	WPI	2002-426119/45.	
XX	Diagnosing	and detecting the progression of liver cancer,	
PT	hepatocellular	carcinoma or metastatic liver tumor in a patient,	
PT	involves	detecting the level of expression of two or more genes in a	
PT	liver	tissue sample	
XX	Claim 1;	SEQ ID NO 3293; 298pp; English.	
XX	The	invention relates to a novel method for diagnosing and detecting the	
CC	progression	of liver cancer, hepatocellular carcinoma or metastatic liver	
CC	tumour	in a patient, and differentiating metastatic liver cancer from	
CC	hepatocellular	carcinoma in a patient, involving detecting the level of	
CC	expression	of two or more genes represented in ABN93503-ABN97455 in a	
CC	tissue	sample. The method of the invention has hepatotropic, and	

CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at pub.wipo.int/pub/published_pct_sequences.

SQ Sequence 2090 BP; 649 A; 384 C; 435 G; 622 T; 0 other;

Query Match 47.8%; Score 782; DB 24; Length 2090;
Best Local Similarity 58.8%; Pred. No. 1.4e-213;
Matches 1108; Conservative 0; Mismatches 490; Indels 12; Gaps 2;

QY	24	ATCATGAGGTC	GACAAAGTCAGCTTGGTATATTCCTGCTCGCAGCTCTCTCTGTG	80
DB	16	ACCAGGATGTC	CTGAAATGGAGCTCAGTCTTCTTCTGCTGATACAGCTCAGTTGTACTTT	75
QY	81	GGCTGTGGATTCT	GTGGAAAGTCCCTGGTGTGCCCTGTGCATGAGCCATTGGCTTAAT	140
DB	76	AGCTCTGGAAGCT	GTGGAAAGTGTAGTGTGCCACAGAAATACACCCATTGGATAAT	135
QY	141	GTC AAGGTCTAAT	CTAGAAGAGCTCATAGTCAGAGGCCATGAGGTAACAGTATTGACTCAC	200
DB	136	ATGAAGACAATCT	GGAAGAGCTGTTCAGAGGGGTCAATGAGGTGACTGTCTCACATCT	195
QY	201	TCAAAGCCTCGTT	AATTTGACTACAGAAAGCCCTCTGCATTGAAATTTGAGGTGGTCCAT	260
DB	196	TCGGCTTCTACT	CTGTGCAATGGCAGTAATCATCTGCTATTAATTTAGAAAGTTTATCCCT	255
QY	261	ATGCC-----	ACAGGACAGAAAGAAATGAAATATTTGTGTGACCTAGCTCTG	311
DB	256	ACATCTTAACTA	AAAAATGATTTGGAAGATTCCTCTGAAATTCCTGATAGATGGATA	315
QY	312	AATGCTCTGCCAG	CGTTATCAACCTGGCAATCAGTTATAAAATTAATGATTTTTTTGTT	371
DB	316	TATGGTGTTTCA	AAAAAATACATTTTGGTCATATTTTTCACAAATTCACAGAAATTCGTGTGG	375
QY	372	GAAATTAAGAGGA	ACTTTAAAAATGATGTGTGAGAGCTTTATCTACAATCAGACGCTTATG	431
DB	376	GAATATTATGACT	ACAGCTACCAAGCTCTGTAAAGATCGAGCTTTTGAATAGAAACTTATG	435
QY	432	AAGAAGCTACAG	AAACCAACTACGATGTAATGCTTATAGACCCGTGATTCCTCGTGGGA	491
DB	436	ATGAAACTACAAG	ATCAAAAGTTTGATGTCTCATTTCTGGCAGATGCCCTTAATCCCTGTGGT	495
QY	492	GACCTGATGCTC	AGTGTGCTTGAGTCCCTTTGTGCTCAGCTTGAATTTCTGTAGGA	551
DB	496	GAGCTACTGGCT	GAACCTATTTACATACCCCTTTCTGTACAGCTCTTCGATCTCTGTGGC	555
QY	552	GGCAATATGAGG	GAAGAGCTGTGGGAACTTCCAGCTCCACTTTCTCTATGTACCTGTGCT	611
DB	556	TACACATTTGAG	AAGAATGCTGGAGGATTTCTGTTCCTCTCTCTATGTACCTGTGTT	615
QY	612	ATCAGAGGACTA	CACAGAGAATGACCTTTCTGGAAAGAGTAAAAATTCATGCTTTCA	671
DB	616	ATGTCAGAAAT	TAAGTCAAAATGATTTTTCATGGAGAGGATAAAAAATATGATACATATG	675
QY	672	GTTTTGTTCAC	TTCTGTGATTCAGGATTCAGACTATCAATTTTGGGAAGAGTTTTATAGT	731
DB	676	CTTTATTTGAC	TTTGGTTTCAATTTATGATCTGAAGAAGTGGGACCACTTTTATAGT	735
QY	732	AAGGCATTAGGA	AGGCCCACTACATTTATGTGAGACTCTGGGAAAGCTGAGATATGGCTA	791
DB	736	GAAGTTCTAGGA	AGACCCACTACATTTATTTGAGACAATGGGAAAGCTGAAATCTGGCTC	795
QY	792	ATACGAAACAT	ATTGGGATTTTGAATTTCCCTCAACCATACCAACCTTACTTTGATTTGTT	851
DB	796	ATTCGAACCTA	TTTGGGATTTTGAATTTTCCCTCGGCCATTTCTTACCAATGTGTGATTTGTT	855
QY	852	GGAGGATTGCAC	TGTAAACCTGCCAAAGCTTTTGCTTAAGAAATGGAAAAATTTTTGTCCAG	911

Db 1186 AATGGCATCTATGAGCGATCTACCATGGGATCCCTATGGTGGGCAATTCCTCTGTTGCG 1245
 QY 1212 GATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAATTC 1271
 Db 1246 GATCAACATGATAAATGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAATTC 1305
 QY 1272 AAAATATGACAAAGCAAGATTTACTGAGGCTTTGAGAACAGTCATTTACCGATTCCTCT 1331
 Db 1306 AGGACCATGCTCAAGTAGAGATTTGCTCAATGCAITGGAAGTCAGTCATTTAATGACCCATC 1365
 QY 1332 TATAAGAGATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTA 1391
 Db 1366 TATAAGAGATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTG 1425
 QY 1392 GATCAGCAGCTCTCTGATGAGTTGTCATGCGCCACAAAGGAGCAGCTGTAGAAATAAATTC 1451
 Db 1426 GATCAGCAGCTCTCTGATGAGTTGTCATGCGCCACAAAGGAGCAGCTGTAGAAATAAATTC 1485
 QY 1452 TCAGCTGCCATGACCTCACCTGGTTCCAGCACTACTCTATGATGATGATGATGATGATGATG 1511
 Db 1486 GTCGAGGCCACAAACCTCACCTGGATCCAGTACCACCTCTTTGGATGATGATGATGATGATG 1545
 QY 1512 CTGACCTGTGGCAACTGCTATATCTTGTTCACAAATGTTTTTATTTTCCTGTCAT 1571
 Db 1546 CTGCGCTGCTGGCAACTGCTATATCTTGTTCACAAATGTTTTTATTTTCCTGTCAT 1605
 QY 1572 AAATTTAATAAACTAGAAAGATAGAAAAGGAGGAATAGATCTTTCCAAA 1621
 Db 1606 AAGCTTGCCAAACAGGAAGAAAGAAAGGATAGTTATATCAAAA 1655

RESULT 10

ABN96789

ID ABN96789 standard; DNA; 2123 BP.

AC ABN96789;

DT 13-AUG-2002 (first entry)

DE Gene #3287 used to diagnose liver cancer.

KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 metastatic liver tumor; cytostatic; expression profile; disease state;
 disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.

PN W0200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer.
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample.

PS Claim 1; SEQ ID NO 3287; 298pp; English.

CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a

CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2123 BP; 676 A; 389 C; 427 G; 631 T; 0 other;

Query Match 46.8%; Score 766.2; DB 24; Length 2123;
 Best Local Similarity 68.4%; Pred. No. 4.9e-209;
 Matches 1078; Conservative 0; Mismatches 489; Indels 9; Gaps 1;

QY 55 TTCTGCTCCTGGAGCTCTTCTGTTGGCTGTGGATTCTGTGGAAAGCTCTGGTGTGGC 114
 Db 8 TGCTGATACAGCTTCAGTTGTTACTTTAGCTCTGGAAGCTGTGAAAGGTGCTAGTGTGGC 67
 QY 115 CCTGTGACATGAGCCATTTGCTTAATGTCGAAGTCAATTTCTAGAGAGCTCATAGTGAGAG 174
 Db 68 CCACAGATACAGCCATTTGATTAATATGAAGACAAATCTGGAAGAGCTTTGTCAGAGG 127
 QY 175 GGCATGAGGTACAGTATTTGACTCACTCAAGCCCTTCGTTAATTTGACTACAGGAAGCCCTT 234
 Db 128 GTCATGAGTGACTGTGTTGACATCTTGGCTTCTACTCTTCTCAATGCCAGTAATCAT 187
 QY 235 CTGCATTTGAAATTTGAGTGGTCCATATGCC-----ACAGGACAGAACAGAGAA 285
 Db 188 CTGCTATTAATTAAGAAGTTTATCTTACATCTTTAACTAAATAATTTTGGAGATTTCTC 247
 QY 286 ATGAATATTTTGTGACCTAGCTCTGAATGTTCTGCCAGGCTTATCAACCTGGCAATCAG 345
 Db 248 TTCTGAAATTTCTCGATAGATGATATATGGTGTTCAAATAATACATTTTGGTCATATT 307
 QY 346 TTATAAATTAATGATTTTTTTTGTGAATAAGAGAACTTTAAATAATGATGTGAGA 405
 Db 308 TTTCACAATTAACAATAATTTGTTGGCATATTTATGACTACAGTAACAAGCTCTGTAAG 367
 QY 405 GCTTTATCTACAATCAGACGCTTATGAAGAAGTACAGAAACCAACTACCATGTAATGC 465
 Db 368 ATGCAGTTTGAATAAGAACTTATGATGAACACTACAAGATCAAGATTTGATGTCAATC 427
 QY 466 TTATAGACCTGTGATTCCTGTGGAGACCTGTAGGTGAGTGTCTTGCAGTCCCTTTTG 525
 Db 428 CTGGAGATCCTGTTTTTCCCTCTGCTGAGCTACTCTCAGAGCTACTTAAACATACCCTTG 487
 QY 526 TGCTCACACTTAGAATTTCTGTAGGAGGCAATATGGAGCAAGCTGTGGGAACCTTCAG 585
 Db 488 TCTACAGGCTCTCGCATCTCTGCTACACANTTGAAGAAGATGGTGGAGGATTTCTGT 547
 QY 586 CTCACATTTCTATGTACCTGTGCTATGACAGGACTTAACAGACAGATGACCTTTCTCG 645
 Db 548 TCCCTCCTCTATGTACCTGTTGTTATGTGTCAGAAATTAATTTGATCAATGATTTTCATGG 607
 QY 646 AAAGAGTAAATAATTCATGCTTTTCAGTTTGTTCACCTCTGATTCAGGATTCAGGATTCAG 705
 Db 608 AGAGGATAAATAATATACATATGCTTTATTTTGTGACTTTTGTGTTTCAAAATTTATGATC 667
 QY 706 ATCAATTTTGGGAAGAGTTTTTATAGTAAGGCATTTAGGAAGCCCACTACATATGTCAGA 765
 Db 668 TGAAGAAGTGGGACCAAGTTTATAGTGAAGTCTTAGGAAGACCCACATACATTTATTGAGA 727
 QY 766 CTGTGGGAAAGAGCTGACATATGCGTAATACGAACATATGGGATTTTGAATTTCCCTCAAC 825
 Db 728 CAATGGGGAAGCTGAAATGTGGCTCATTTGGAACCTATTTGGGATTTTGAATTTCCCTCGCC 787
 QY 826 CATACCAACCTAACTTTGAGTTTGTGGAGGATTTGCACTGTAACCTGCAACAGCTTTGCG 885
 Db 788 CATCTTTACCAATTTGATTTTGTGGAGGACTTCACTGTAACCAAGCAACCCCTGCG 847

```
Qy 886 CTAAGGAAATGGAAATATTTGTCCAGAGTTCAGGGGAAGATGGTATTGTGGTGTCTTC 945
Db 848 CTAAGGAAATGGAAAGATTTGTCCAGAGTTCAGGGGAAGATGGTATTGTGGTGTCTTC 907
Qy 946 TGGGGTCACTCTTTCAAAATGTTACAGAGAAAGGCTTAATATCATTCCTTCAGCCCTTG 1005
Db 908 TGGGGTCACTCTTTCAAAATGTTACAGAGAAAGGCTTAATATCATTCCTTCAGCCCTTG 967
Qy 1006 CCCAGATCCCACAGAGAGTGTATGGAGGTACAAAGGAAAGAAACCACCATAGGAG 1065
Db 968 CCCAGATCCCACAGAGAGTGTATGGAGGTACAAAGGAAAGAAACCACCATAGGAG 1027
Qy 1066 CCAATACCTGGCTGTATGATGATACCCAGAAATGATCTTCTGGTATCCCAAAACCA 1125
Db 1028 CCAATACCTGGCTGTATGATGATACCCAGAAATGATCTTCTGGTATCCCAAAACCA 1087
Qy 1126 AAGCTTTTATCATTGATGGTGAATGAATGGGATCTATGAAGCTATTACCATGGGTCC 1185
Db 1088 AAGCTTTTATCATTGATGGTGAATGAATGGGATCTATGAAGCTATTACCATGGGTCC 1147
Qy 1186 CTATGGTGGGAGTCCCATATTTGGTGTATCAGCTTGATTAACATAGCTCACATGAAGGCA 1245
Db 1148 CTATGGTGGGAGTCCCATATTTGGTGTATCAGCTTGATTAACATAGCTCACATGAAGGCA 1207
Qy 1246 AAGGAGCAGCTGTAGAATAAACTTCAAACTATGACAAGCGAAGATTTACTGAGGCTT 1305
Db 1208 AAGGAGCAGCTGTAGAATAAACTTCAAACTATGACAAGCGAAGATTTACTGAGGCTT 1267
Qy 1306 TGACACAGCTATACCGATTCCTTTATTAAGAGAGATGCTATCAGATTTATCAAGATTC 1365
Db 1268 TGAAGTCACTATTAATGACCCAGCTCTATAAAGAGATGCTATGAAATTTATCAAGATTC 1327
Qy 1366 ACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGCTTCTGGATCGAGTTTGTCTATGC 1425
Db 1328 ATCATGACCAACCAATGAGCCCTGATCGAGCAGCTTCTGGATCGAGTTTGTCTATGC 1387
Qy 1426 GCCACAAGAGCCAAAGCAGCTGGATCAGCTGCCCCATGACCTCACCTGGTTCACGACCT 1485
Db 1388 GCCACAAGAGCCAAAGCAGCTGGATCAGCTGCCCCATGACCTCACCTGGTTCACGACCT 1447
Qy 1486 ACTATATAGATGATGGGTTCCTGCTGACCTGCTGGCAACCTGCTATATTTCTGTTC 1545
Db 1448 ACTCTTGGATGTATGATGATCTCTGCTGGCTGGGCGCAACTGTGATATTTATCATCA 1507
Qy 1546 CAAATGTTTTTTTATTTCTGTCAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGG 1605
Db 1508 CAAATTTTGGCTGTTTGTTCGAAAGCTTGCCAAAAGAGAAAGAAAGAAAGAG 1567
Qy 1606 AATGATCTTTTCCAAA 1621
Db 1568 ATTAGTTATATCAAAA 1583
```

RESULT 11

```
AA295199
ID AA295199 standard; DNA; 2092 BP.
XX
AC AA295199;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human UDP-glucuronosyltransferase 2B4 nucleotide sequence.
XX
KW UDP-glucuronosyltransferase 2B4; UGT2B4; polymorphism; metabolism; SNPs;
KW drug interaction; detect; human; single nucleotide polymorphism; ds.
XX
OS Homo sapiens.
XX
PN WO200006776-A1.
XX
PD 10-FEB-2000.
XX
PF 22-JUL-1999; 99WO-US16675.
```

```
XX 28-JUL-1998; 98US-0094391.
XX (AXYS-) AXYS PHARM INC.
XX Galvin M, Miller A, Penny L, Riedy M;
XX WPI; 2000-195321/17.
XX P-PSDB: AAY78933.
XX Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
XX genotyping individuals to predict rate of metabolism of substrates and
XX for identifying potential drug interactions
XX PS Disclosure; Page 34-36; 72pp; English.
XX
XX This sequence represents the human UDP-glucuronosyltransferase 2B4
XX (UGT2B4) gene. UDP-glucuronosyltransferase (UGTs) are a family of
XX enzymes that catalyze the glucuronic acid conjugation of a wide range of
XX endogenous and exogenous substrates. The UGT2B gene subfamily encode
XX steroid metabolizing isoforms in the liver. Alteration of the expression
XX or function of UGTs may effect drug metabolism. The invention relates to
XX non-chromosomal nucleic acid molecules, which comprise human UGT2B
XX sequence polymorphisms (see AA295051-295110). Probes which detect the
XX UGT2B locus polymorphisms can be used to detect altered UGT2B metabolism
XX of a substrate in an individual. The nucleic acid molecules comprising a
XX human UGT2B sequence polymorphism can be used in screening assays for
XX genotyping individuals, also to predict their rate of metabolism of
XX UGT2B substrate, potential drug-drug interactions and adverse side
XX effects. The polymorphisms can be used as single nucleotide polymorphisms
XX (SNPs) for detecting genetic linkage related to phenotypic variation in
XX activity or expression of UGT2B protein. The polymorphism containing
XX nucleic acid molecules may also be used for generating genetically
XX modified non-human animals and for obtaining site specific gene
XX modification in cell lines.
XX
XX Sequence 2092 BP; 639 A; 398 C; 438 G; 617 T; 0 other;
```

```
Query Match 46.8%; Score 766; DB 21; Length 2092;
Best Local Similarity 68.7%; Pred. No. 5.5e-209;
Matches 1102; Conservative 0; Mismatches 490; Indels 12; Gaps 3;

Qy 23 CATCATGAGGCTGACAAAGTCAGCTTGGTATATTTCTGCTCTCGACGCTCTCTGT 79
Db 31 CATCAGGATGCTATGAAATGGACTTCAGCTCTTCTGCTATACAGCTGAGCTGTACTT 90
Qy 80 TGGCTGTGGATTCGTGGGAAGTCCTGGTGTGGCCCTGTGCATGACATGACCAATGCTTAA 139
Db 91 TAGCTCTGGGAGTTGTGGAAAGGTGCTGGTGTGGCCACAGAAATTCAGCCACTGGATGAA 150
Qy 140 TGTCAGGCTCATTTCTAGAAGAGCTCATAGTCAGAGGGCCATGAGGTAACAGATTGACTCA 199
Db 151 TATAAGACAACTCTGGATGAACCTGTCACAGAGGTCATCAGGTGACTGTATTGGCATC 210
Qy 200 CTCAAAGCCTTCGTTAATTTGACTACAGGAAGCCTTCTGCATTTGAAATTTGAGTGGTCCA 259
Db 211 TTCAGCTTCCATTTCTTTCGATCCCAAGCCATCTACTCTTAAATTTGAAGTTATCC 270
Qy 260 TATGCCACAGGACAGAAAGAAATGAATATTTGTCACCTAGCTCTGA-----A 313
Db 271 TGTATCTTTAACTAAACTGAGTTTGGAGATATTTATCAGCAGCTGGTGTAGAGATGGCC 330
Qy 314 TGTCTTGCCAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATATGATTTTTTTGTGTA 373
Db 331 AGAACTTCCAAAGACACATTTTGGTCATATTTTTCACAACTACAGAAATCATGTGGAC 390
Qy 374 AATAAGAGGAACCTTAAATAATGATGTGTGAGAGCTTTTATCTACAAATCAGAGCCTTATGAA 433
Db 391 ATTTAATGACATACTTAGAAAGTCTCTGAAGGATATAGTTTCAATTAAGAAACTTATGAA 450
Qy 434 GAAGCTACAGGAAACCAACTACGATGTAATCTTTATGACCTCTGTATTCCTCTGGGAGA 493
Db 451 GAAACTACAGGAGTCAAGATTTGATGTTCTTTCAGAGATCTGTTTCCCTCTGGTGA 510
```


is at risk for breast cancer or has breast cancer. BCSG antibodies
CC labelled with paramagnetic ions or radioisotopes is useful for imaging
CC breast cancers, while BCSG antibodies conjugated to a cytotoxic agent is
CC useful for treating breast cancer. BCSG is useful in the rational design
CC of new therapeutics for imaging and treating cancers. BCSG is also used
CC in gene therapy. The present DNA sequence is human breast cancer specific
XX gene-2 (BCSG-2) or Gene ID 480489.

Sequence 1722 BP; 523 A; 319 C; 371 G; 509 T; 0 other;

Query Match 46.1%; Score 754.8; DB 22; Length 1722;
Best Local Similarity 67.8%; Pred. No. 8.1e-206;
Matches 1088; Conservative 0; Mismatches 507; Indels 9; Gaps 2;

QY 23 CATCATGAGGTCTGACAACTGACGTTTGGTATTTCTGCTCCGACGCTCTCTGT---GT 79
DB 3 CACCAGGATGACTCTGAATGGACTTCAGTTCTTCTGCTGATACATCTCAGTTGTACTT 62
QY 80 TGGCTGTGGATCTGTGGGAAAGTCTGTGTGGCCCTGTGACATGAGCCATTGGCTTAA 139
DB 63 TAGCTCTGGAGTTGTGGAAAGTCTGTGTGGCCGCGAGAATACAGCCATTGGATGAA 122
QY 140 TGTCAAGGTCATCTAGAGAGCTCATCTGAGAGCCCATGAGTAACAGTATTGACTCA 199
DB 123 TATGAAGCAATCCTGAAGAGCTTGTTCAGAGAGGTCATGAGGTGACTGTACTGGCATC 182
QY 200 CTCAAAGCCTTCGTTAATTGACTACAGGAAGCCTTCTGCATTTGAAATTTGAGTGTGCA 259
DB 183 TTCAGCTTCCATCTTTTGTATGCCCAATGATGCATCCACTCTTAAATTTGAAGTTATCC 242
QY 260 TATGCCACAGGACAGAACAGAAATGAATATTTTGTGACCTAGCTCTGA-----A 313
DB 243 TACATCTTTAACTAAAACTGAAATTCAGAAATATCATCATGCAACAGGTTAAGAGATGGTC 302
QY 314 TGTCTTCCAGGCTTATCAACCTGGCAATCAGTTATATAATTTAAATGATTTTGTGTTGA 373
DB 303 AGACATTCGAAAGATAGCTTTTGGTTATATTTTCCAAAGAACAAAGAAATCCTGTGGGA 362
QY 374 AATAAGAGAACTTTAAAAATGATGTGAGAGCTTTATCTCAACTCAGACGCTTATGAA 433
DB 363 ATTATATGACATATTTAGAACTTCTGTAAGATGTAGTTTCAATTAAGAAAGTTATGAA 422
QY 434 GAAGCTACAGGAAACCACTACGATGTAAGTATAGACCCCTGTGATTCCTCTGTGGAGA 493
DB 423 AAAACTACAGAGTCAAGATTTGACATCGTTTTTGGCAGATGCTGTTTTTCCCTGTGGTGA 482
QY 494 CTTGATGGCTGAGTCTGCTTCCAGTCCCTTTTGTGCTCACACTTAGAATTTCTAGAGG 553
DB 483 GCTGCTGGCTGCTACTTAACATACGCTTGTGTACAGTCTCCGCTTTACTCCTGGCTA 542
QY 554 CAATATGGAGCGAAGCTGTGGGAACTTCCAGCTCCACTTTCTATGTACCTGTGCTAT 613
DB 543 CACAATTTGAAGGCACAGTGGAGGACTGATTTCCCTCCTCTCTACATACCTATTTGTTAT 602
QY 614 GACAGACTTAACAGACAGTAACCTTTCTGGAAGAGGTAAAAATTCATGCTTTTCACT 673
DB 603 GTCAAAATTAAGTGATCAATGACTTTTCATGGAGAGGGTAAAAAATATGATCTATGCT 662
QY 674 TTTGTTCCACTTCTGATTCAGGATTTACGACTATCATATTTTGGGAGAGTTTATAGTAA 733
DB 663 TTATTTTGACTTTTGTGTTCCAAATGTCTGATATGAAGAAGTGGGATCAGTTTACASTGA 722
QY 734 GGCATTAGGAAGGCCCACTACATTTATGTAGAGCTGTGGGAAAGCTGAGATATGGCTAAT 793
DB 723 AGTTTTAGGAGACCCACTACCTTATTTGAGCAATGGGAAAGCTGACATATGGCTTAT 782
QY 794 ACGAATATATGGGATTTTGAATTTCTCAACCATACCAACTTAACCTTTGAGTTTGTGG 853
DB 783 GCGAAACTCTCGAGTTTCAATTTCTCATCATCTTTACCAAAAGCTGATTTTGTG 842
QY 854 AGGATTTGCACTGTAAACCTGCCAAAGCTTTGCTTAAGGAAATGGGAAATTTGTCCAGAG 913
DB 843 AGGATTTCCACTGCAAACTGCCAAAGCTTACCTAACGAAATGGAGAGTTTGTGTACAGAG 902

QY 914 TTCAGGGAAGATGGTATTGTGGTGTCTTCTGGGTACACTGTTTCAAAATGTTACAGA 973
DB 903 CTCTGGAGAAATGGTGTGGTGTCTTCTGGGTACAGTAAATGATTAACATGACAGC 962
QY 974 AGAAAAGGCTAATATCATTTGCTTCAGCCCTTCCAGAGATCCACAGAGGTGTTATGGAG 1033
DB 963 AGAAAGGGCCAATGTAATTTGCAACAGACCCCTTGCCAAAGATCCCAAAAGGTTCTGTGGAG 1022
QY 1034 GTACAAAGGAAAGAAACCATCCACATTTAGGACCCATCTCGGCTGTATGATTGGATACC 1093
DB 1023 ATTTGACGGGAATAAACAGATGCCCTTAGGTTCTCAATCTCGGCTGTCAAGTGGATACC 1082
QY 1094 CCAGAAATGATCTTCTTGGTTCATCCCAAAACCAAGGCTTTTATCACTCATCTGGTGGAAATGAA 1153
DB 1083 CCAGAAATGACCTTCTAGGTCATCCAAAAACAGAGCTTTTATACTCATCTGGTGGACCAA 1142
QY 1154 TGGGATCTATGAAGCTATTTTACATGGGGTCCCTATGTTGGGAGTTCCTCATATTTGGTGA 1213
DB 1143 TGGCATCTATGAGGCAATCTAGCATGGGATCCCTATGTTGGCATTTCCATTTGTTTGA 1202
QY 1214 TCAGCTTGATACATAGCTCACATGAAGCCCAAGGAGCAGCTGTAGAATTAACCTTCAA 1273
DB 1203 TCAACCTGATAACATTTGCTCACATGAAGCCCAAGGAGCAGCTGTAGATTGGACTTCAA 1262
QY 1274 AACTATGACAAAGCAAGATTTACTGAGGGCTTTTGAGAACAGCTATTACCGATTTCCTCTTA 1333
DB 1263 CACAATGTCGAGTACAGACCTGCTGATGCTCACTGAAGACAGTAATTAATGATCCTTTATA 1322
QY 1334 TAAAGAAATGCTATGAGATTTATCAAGAAATCAACATGATCAACCTGTAAAGCCCTAGA 1393
DB 1323 TAAAGAAATATATGAAATTTATCAAGAAATCAACATGATCAACCAAGTAAAGCCCTGGA 1382
QY 1394 TCGACGACTTCTGGATCGAGTTTGTCTATGCCCAAGGAGCAGCAACGACCTCGGATC 1453
DB 1383 TCGAGCAGTCTTCTGGATTTGTAATTTGTATGCCCAAGGAGCAGCAACGACCTCGGAT 1442
QY 1454 AGCTGCCCATGACCTCACCTGGTTCAGACACTACTCTATAGATGTGATTGGGTTCCTGCT 1513
DB 1443 TGCAGCCCATGACCTCACCTGGTTCAGTACCACTCTTTGGATGTGATTGGGTTCCTGCT 1502
QY 1514 GACCTGTGGGCAACTGCTATATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1573
DB 1503 GGCCTGTGTGGCAACTGTGATATTTATCATCAAAAGTTTTGTCTGTTTGTGTTGTTGTTGTT 1562
QY 1574 ATTTAATAAACTAGAAAGATAGAAAGGAGGAATAGATCTTTC 1617
DB 1563 GTTTGCTAGAAAGGGAAGAGGAAAAAGAGATTAGTTATGTC 1606

RESULT 14

ABN95622

ID ABN95622 standard; DNA; 1855 BP.

XX

AC ABN95622;

XX

DT 13-AUG-2002 (first entry)

XX

DE Gene #2120 used to diagnose liver cancer.

XX

KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX

OS Homo sapiens.

XX

PN WO200229103-A2.

XX

PD 11-APR-2002.

XX

XX 02-OCT-2001; 2001WO-US30589.

XX

XX 02-OCT-2000; 2000US-237054P.

XX

PR

xx (GENE-) GENE LOGIC INC.
xx Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
xx WPI; 2002-426119/45.
xx
xx Diagnosing and detecting the progression of liver cancer,
xx hepatocellular carcinoma or metastatic liver tumor in a patient,
xx involves detecting the level of expression of two or more genes in a
xx liver tissue sample
xx
xx Claim 1: SEQ ID NO 2120; 298pp; English.
xx
xx The invention relates to a novel method for diagnosing and detecting the
xx progression of liver cancer, hepatocellular carcinoma or metastatic liver
xx tumour in a patient, and differentiating metastatic liver cancer from
xx hepatocellular carcinoma in a patient, involving detecting the level of
xx expression of two or more genes represented in ABN93503-ABN97455 in a
xx tissue sample. The method of the invention has hepatotropic, and
xx cytostatic activity. The method is useful for diagnosing and detecting
xx the progression of liver cancer, hepatocellular carcinoma and metastatic
xx liver carcinoma in a patient. The method is useful for identifying
xx expression profiles which serve as useful diagnostic markers as well as
xx markers that can be used to monitor disease states, disease progression,
xx drug toxicity, drug efficacy and drug metabolism.
xx Note: The sequence data for this patent did not form part of the printed
xx specification, but was obtained in electronic format directly from WIPO
xx at ftp.wipo.int/pub/published_pct_sequences.

xx SQ Sequence 1855 BP; 571 A; 337 C; 394 G; 553 T; 0 other;

Query Match 46.1%; Score 754.8; DB 24; Length 1855;
Best Local Similarity 67.8%; Pred. No. 8.5e-206;
Matches 1088; Conservative 0; Mismatches 507; Indels 9; Gaps 2;

QY 23 CATCATGAGCTCGACAAAGTCAGCTTTGGTATTTCTGCTCCTCGAGCT----CTTCTGTGT 79
DB 8 CACCAGGATGCTGTCGAAATGGAGCTTCAGTAAATTTTGTCTAATCAACTGAGCTTTTGCTT 67
QY 80 TGCGTGTGGATTCTGTGGGAAAGTCCTGTGGTGGCCCTGTGACATGAGCCATGGCTTAA 139
DB 68 TAGCTCTGGGAATTTGGGAAAGTGTGGTGGTGGGAGCAGATACAGCCATTTGGATGAA 127
QY 140 TGTCAAGGCTATCTAGAAGAGCTCATGATGAGAGGCCATGAGTACACATTTGACTCA 199
DB 128 TATAAGACAACTCCTGGATGAGCTTATTCAGAGAGGTCATGAGTGTACTGTGGCATC 187
QY 200 CTCAAGCCCTTGTAAATGACTACAGGAGCCCTTGTGATTTGAAATTTGAGGTGCTCA 259
DB 188 TTCAGCTTCCATTTCTTTTGGATCCCAACTCATCGCTTTTAAATTTGAAATTTATCC 247
QY 260 TATGCCACAGGACAGACAGAAATGAAATATTGTTGACCTAGCTCTGA-----A 313
DB 248 CACATCTTTAACTAAACTGAGTTGGAGAAATTCATCATCCACAGATTAAGAGATGGTC 307
QY 314 TGTCTTGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATGATTTTGTGTA 373
DB 308 AGACCTTCCAAAGAGATACATTTTGGTTATATTTTTCACAGAGTACAGGAAATCATGTCAAT 367
QY 374 AATAAGAGGAACCTTAAATATGATGTGAGAGCTTTATCTACAAATCAGAGCTTATGAA 433
DB 368 ATTTGGTGACATAACTAGAAAGTTCTGTAAGATGTAGTTTCAAAATGAAGAAATTTATGAA 427
QY 434 GAAGCTACAGGAACCAACTACCATGTAATGCTTATAGACCCGTGATTCCTGTGGAGA 493
DB 428 AAAAGTACAAGAGTCAAGATTTGACGTGATTTTGGCAGATGCTATTTTCCCTGTAGTGA 487
QY 494 CCTGATGGCTGAGTTGCTTCCAGTCCCTTTTGTGCTCACACATTTAGAAATTTCTGTAGGAGG 553
DB 488 GCTGCTGGCTGAGCTATTTAAACATACCCCTTTGTGTACAGTCTCAGCTTCTCTCTCTGCTA 547
QY 554 CAATATGGAGCGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCTATGTACCTGTGCTAT 613

RESULT 15

DB 548 CACTTTTGAAGACATAGTGGAGGATTTATTTCCCTCTCTCTACCTGTTGTTAT 607
QY 614 GACAGGACTAACACAGAGAAATGACCTTTCTGGGAAGAGATAAAAATTCATGCTTTCAGT 673
DB 608 GTCAGAAATTAACATGATCAAAATGACCTTTCATGGAGAGGGTAAAAAATATGATCTATGCT 667
QY 674 TTTGTTCCACTTCTGGATTTCAGGATTCAGGACTATCATTTTGGGAAGAGTTTATAGTAA 733
DB 668 TTACTTTGACTTTTGGTTCGAAATATTTGACATGAAGAAGTGGGATCAGTTTATAGTGA 727
QY 734 GGCATTAGGAAGCCCACTACATATTTATGTAGACTGTGGGAAAGCTGAGATATGGCTAAT 793
DB 728 AGTTCTAGGAAGACCCCACTACGTTATCTGAGACAAATGGGAAAGCTGACCTATGCTTAT 787
QY 794 ACNACATATTTGGATTTTCAATTTCTCAACCATACCAACCTAACCTTTCAGTTTGTGG 853
DB 788 TCAGAACTCTCGAAATTTTTCAGTTTCTCATCACTCTTACCAANTGTTGATTTGTGG 847
QY 854 AGGATTGCATCTGTAACCTGCCAAAGCTTTGCCCTAAGGAAATGGAATTTTGTCCAGAG 913
DB 848 AGGACTCCACTGCAAACTGCCAAACCCCTGCCCTAAGGAAATGGAAGACTTTGTACAGAG 907
QY 914 TTCAAGGGAAGATGTTATTTGTTGTTTCTCTGGGGTCACTGTTTCAAAATGTTACAGA 973
DB 908 CTCTGGAGAAATGTTGTTGTTTCTCTGGGGTCAATGGTCACTAGTAAACATGACAGA 967
QY 974 AGAAAGGCTTAATATCATTTCTCAGCCCTTGCACAGATCCCAAGAGTGTGTTATGAG 1033
DB 968 AGAAGGGCCCAACGTAATTCATCAGCCCTGGCCAGATCCCAAGAGTTCGTGAGAG 1027
QY 1034 GTACAAAGGAAAAAACCATCCATTTAGAGGCCAATPACTCGGCTGTATGATGGATACC 1093
DB 1028 ATTTGATGGGAATAAACAGATACCTTAGTCTCAATACTCGGCTGTATAGTGGATACC 1087
QY 1094 CCAAGATGATCTTCTTGGTTCATCCCAAAACCAAGACTTTTATCACTCATGTTGGAATGAA 1153
DB 1088 CCAGAAATGACCTTCTTAGTTCATCCAAAGACAGAGCTTTTATACTCATGTTGGAGCCAA 1147
QY 1154 TGGGATCTATGAAGCTATTTACCATGGGTGCCCTATGTTGGGAGTTCCTCATATTTGTTGA 1213
DB 1148 TGGCATCTACGAGGCAATCTACCATGGATCCCTATGTTGGGATTCCTATTTGCGGA 1207
QY 1214 TCAGCTTGAATACATAGCTCAGATGAGGCGCAAGAGAGAGCTGTGAATAATAAATCTCAA 1273
DB 1208 TCAACCTGATAAACATTTGTCACATGAGGCGGAGAGAGCTGTTAGAGTGGACTTCAA 1267
QY 1274 AACTATCAGAGCGAAGATTTTACTGAGGCTTTGAGAACACATCATTTACCGATTCCTCTTA 1333
DB 1268 CACAATGTCGAGTACAGACTTGTGATGATTTGAAGAGAGTAAATTAATGATCTTCATA 1327
QY 1334 TAAAGAGAAATGCTATGAGATTATCAAGAAATTCACCATGATCAACCTGTAAGCCCCCTAGA 1393
DB 1328 TAAAGAGAAATGTTATGAATTTATCAAGAAATTCACCATGATCAACCTGTAAGCCCCCTAGA 1387
QY 1394 TCAGAGAGTCTTCTGGATCAGAGTTTGTATGCGGCCCAAAAGGAGCCAGACCTGCGCATC 1453
DB 1388 TCAGAGAGTCTTCTGGATTTGAAATTTGTCATGCGCCACAAAGAGAGCTTAAACCTTCGGGT 1447
QY 1454 AGCTGCCATGAGCTCACTACCTGTTTCCAGACTACTCTATAGATGTGATGGTTCCCTGCT 1513
DB 1448 TGCAGCCCAAGAGCTCACTGTTTCCAGTACCCTCTTTGGATGTGATTTGGTTCCCTGCT 1507
QY 1514 GACTGTGTGGCAACTGCTATATTTCTTGTTCACAAAATGTTTTTATTTTCTCTCAAAA 1573
DB 1508 GGTCTGTGTGGCAACTGCTGATATTTATGCTCACAATAATGTTGCTGTGTTGTTTCTGAA 1567
QY 1574 ATTTAATAAAGCTAGAAAGATAGAAAGAGAGGGAATAGATCTTTC 1617
DB 1568 GTTTGCTAGAAAGCAAGGAAAGAGGGAATAATGATTAGTTATATC 1611

Qy	674	TTTGTGCCACTCTCGGATTCAGGATTCAGCAATATCATATTTTGGGAAGAGCTTTTATATAGTAA	733
Db	668	TTACTTTGACTTTTGGTTCGAAATATTTGACATGAAGAAGTGGGATCAGTTTTATATAGTGA	727
Qy	734	GGCATTAGGAAGGCCACATACATATATGTGAGACTGTGGGAAGAACTGAGATATGGCCTAAT	793
Db	728	AGTTCTAGGAAGACCCACTACGTTATCTGAGACAATGGGGAAGAGCTGACGATATGGCTTAT	787
Qy	794	ACGAACATATTTGGGATTTTGAATTTCCCTCAACCATACCAACCTTAACCTTTGAGTTTCTTGG	853
Db	788	TCGAAACTCCTGGAAATTTTTCAGTTTCTCATCCACTCTTACCAAAATGTTGATTTTCTTGG	847
Qy	854	AGGATTCGACTGTAAACCTGCCAAAGCTTGGCCCTAAGGAATGGAAATTTTGTCCACAG	913
Db	848	AGGACTCCACTCGAAACCTGCCAAACCCCTGCCCTAAGGAATGGAAAGACTTTGTACAGAG	907
Qy	914	TTCCAGGGAAGATGATATCTGTGTGTTTCTCTCTGGGTCACTCTTTCAAAATGTTACAGA	973
Db	908	CTCTGGNAAAAATGGTGTGTGTGTTTCTCTCTGGGTCAATGGTCACTGATGATGACAGA	967
Qy	974	AGAAAAGGCTTAATATCATCTGCTTCAGCCCTTGCCAGATCCCAAGAGTGTATGGAG	1033
Db	968	AGAAAGGCCCAACGTAATTCGATCAGCCCTGGCCAGATCCCAACAAAGGTTCTGTGGAG	1027
Qy	1034	GTACAAAGGAAAAAACCATCCACATTAGAGAGCCAATACTCGGCTGTATGATTTGGATACC	1093
Db	1028	ATTGTATGGGAATAAACCCAGATACCTTTAGGTCTCAATCTCTCGGCTGTATAAAGTGGATACC	1087
Qy	1094	CCAGATGATCTTCTTGGTCATCCCAAAACCAAGCTTTTATCACCTCATGGTGGGAATGAA	1153
Db	1088	CCAGATGACCTTCTAGGTCAATCCAAAGACCAGAGCTTTTATAACTCATGGTGGAGCCAA	1147
Qy	1154	TGGGATCTATGAAGCTATTTACCATTGGGTGCCCTATGGTGGAGTTCCTCCCATATTTGGTGA	1213
Db	1148	TGGCATCTACGAGGCAATCTACCATGGGATCCCTATGGTGGGATTCCTATGTTGCCGA	1207
Qy	1214	TCAGCTGTATACATAGCTACATGAAGGCCAAAGAGCAGCTGTAGAAAATAAATCTCAA	1273
Db	1208	TCAACCTGTATTAATCTGCTCACATGAAGGCCAGGGAGCAGCTGTAGAGTGGACTCAA	1267
Qy	1274	AACTATGACAAGCGAAGATTTACTGAGGGCTTTTGAGAACAGTCAATACCGATTCTCTTA	1333
Db	1268	CACATGTCAGTACAGACTTGTGATGCAATTTGAAGAGAGTAATTAATGATCCTTCATA	1327
Qy	1334	TAAAGAGATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTGTAAAGCCCCTAGA	1393
Db	1328	TAAAGAGATGTTATGAAATTTATCAAGAATTCAAACATGATCAACCACTGAAGCCCCCTGGA	1387
Qy	1394	TCGAGCAGTCTTCTGGATCGAGTTTGTCAATGCCCCACAAGAGGCCAACGACCTGGCATC	1453
Db	1388	TCGAGCAGTCTTCTGGATGAAATTTGTCAATGCCCCACAAGAGAGCTTAACACCTTCGGGT	1447
Qy	1454	AGCTGCCCATGACTCACCTGGTTCACGCACTACTCTATAGATGTGATTTGGGTTCCTGCT	1513
Db	1448	TGACGCCACGACCTCACCTGGTTCAGTACCACCTCTTTGGATGTGATTTGGGTTCCTGCT	1507
Qy	1514	GACCTGTGTGGCAACTGCTATCTTCTTGTTCACAAAAATGTTTTTATTTTCTGTGTCAAAA	1573
Db	1508	GGTCTGTGTGGCAACTGTGTATTTTATTCGTCACAAAATGTTGTCTGTTTGTGTTCTCGAA	1567
Qy	1574	ATTTAATAAATAGAAAGATAGAAAAGAGGGGAATAGATCTTTC	1617
Db	1568	TTTTGCTAGAAAAGCAAGAGGGAAGGAATGATTTAGTTATATC	1611

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 08:51:42 ; Search time 1786.01 Seconds
(without alignments)

14835.236 Million cell updates/sec

Title:

US-09-981-353-165

Perfect score:

1636

Sequence:

1 gatcagtggtgagggaact.....ccaaattcaagaagacctg 1636

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	882.8	54.0	2099	11	AK008601 Mus muscu
2	698.8	42.7	2573	11	AK004971 Mus muscu
3	625.2	38.2	960	14	BQ713091 AGENCOURT
4	623.2	38.1	1004	14	BQ942104 AGENCOURT
5	622.4	38.0	1896	11	AK002736 Mus muscu
6	600	36.7	927	14	BQ925596 AGENCOURT

7	552.4	33.8	930	14	BQ934372
8	520.4	31.8	1030	14	BQ714919
9	492.6	30.1	913	14	BQ714368
10	477	29.2	823	12	BF680999
11	467.4	28.6	671	12	BG249053
12	464.4	28.4	977	13	B1331895
13	464.4	28.4	871	14	BQ900474
14	463.4	28.3	938	14	BQ942682
15	456.4	27.9	891	12	BG189133
16	455.6	27.8	770	12	BG201991
17	435	26.6	746	12	BF688309
18	424.8	26.0	996	14	BQ714776
19	418.6	25.6	801	12	BG212458
20	399.6	24.4	759	13	B1330877
21	399.4	24.4	554	10	BE013295
22	398	24.3	962	13	BG969620
23	394	24.1	728	12	BG569536
24	394	24.1	761	9	A1118428
25	387.2	23.7	1005	12	BG167171
26	383	23.4	1192	12	BF789664
27	378.4	23.1	886	12	BG402016
28	373	22.8	651	9	AA255327
29	369.2	22.6	830	12	BG195984
30	367.2	22.4	836	9	A1529553
31	366.2	22.4	935	12	BG427877
32	364.2	22.3	812	13	B1221130
33	361.4	22.1	823	12	BG193837
34	351.8	21.5	882	13	B1144244
35	349.6	21.4	906	9	A1529750
36	342.8	21.0	814	12	BG428781
37	340.4	20.8	513	12	BF105272
38	337.8	20.6	1022	14	BM925650
39	334.2	20.4	696	13	BG971909
40	332.6	20.3	2450	11	BC011409
41	331.4	20.3	557	12	BF080461
42	331	20.2	639	12	BF322562
43	330.2	20.2	934	12	BF786200
44	328.6	20.1	2202	11	AK002629
45	321.6	19.7	892	12	BG208263

ALIGNMENTS

RESULT 1	AK008601	2099 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010321J07:homolog to UDP				
DEFINITION	GLUCURONOSYLTRANSFERASE UGT2A3, full insert sequence.				
ACCESSION	AK008601				
VERSION	AK008601.1	GI:12842884			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male small intestine cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	clone:2010321J07.				
	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, S., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, F., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 (bases 1 to 2099)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Onno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, F., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Location/Qualifiers
1. .2099
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2010321J07"
/db_xref="MGI:1904826"
/db_xref="taxon:10090"
/clone="2010321J07"
/sex="male"
/tissue_type="small intestine"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

COMMENT

Location/Qualifiers
1. .2099
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2010321J07"
/db_xref="MGI:1904826"
/db_xref="taxon:10090"
/clone="2010321J07"
/sex="male"
/tissue_type="small intestine"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

FEATURES
source

Location/Qualifiers
1. .2099
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2010321J07"
/db_xref="MGI:1904826"
/db_xref="taxon:10090"
/clone="2010321J07"
/sex="male"
/tissue_type="small intestine"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

CDS

/dev_stage="adult"
15. .1619
/notes="data source:SPTR, source key:Q9R110, evidence:ISS homolog to UDP GLUCURONOSYLTRANSFERASE UGT2A3 putative"
/codon_start=1
/protein_id="BAB25770.1"
/db_xref="GI:12842885"
/db_xref="MGI:1919344"
/translation="MYSEKVAFFELLOLCWAGCGFCSKVLVWPCDMSHLNKLKLTLE ELGARGHEVTLKYPSTIIIDSKRIPLHFNIPILLYEIEFAENLAEINLANVAVIPN LSLUAAKTLQDFQVGTDFESICRSVLYNQKFMKLRDAQDVVVIDVPPVCGELY AEVLQIPFVYTLRFSGMYMEKHCQGLPIPLSVPPVSELTDMNTTDEKRVKMFESL LFGLWQKQDFADFDYFSETLGRPTTFCYTGADWLIRTYWDFEPPYLPNPEF VLVWYSGKPAIILGNTNFWIPONDLLGHPKTKAFITHTGNTGVEAYLHGVPMWG VPMUGDPNIAHMEAKGALKYSISTMTSDLLSAVRVINEPSYKENARLSRIHH DQVPLDRAVFWIEFMRHKGAKHLRAAHLDSWFIQHSLDVIGFLLCVLTFTFII TRFCLFVQKLYMKESKMGNRKKN"
BASE COUNT 618 a 406 c 467 g 608 t
ORIGIN
Query Match 54.0%; Score 882.8; DB 11; Length 2099;
Best Local Similarity 73.4%; Pred. No. 8.4e-204;
Matches 1173; Conservative 0; Mismatches 412; Indels 13; Gaps 3;
QY 23 CATCATGAGCTCTGACAAAGTCAGCTTTGGTATTCTGCTCTGCAGCTCTTCTGTGTGG 82
DB 11 CAATATGGTCTCTGAAAATATGTTCGGGCATTTTCTGCTGCAGCTTTGCTGGCGG 70
QY 83 CTGTGATTCGTGGGAAGTCCCTGGTGGCCCTGTGACATGAGCATTGGCTTAATGT 142
DB 71 CTGTGATTCGTGAGCAAGCTCCTGCTGGCCCTGTGATATGAGCCATGGCTGAATCT 130
QY 143 CAAGGTCATTCAGAGAGCTCATAGTGAGAGCCATGAGTAACTATTTGACTCACTC 202
DB 131 AAAGACTATTCTTGGAGGAGCTTGGAGCAAGGCGGACAGGTACAGCTCTGAAATACC 190
QY 203 AAAGCCTTCCTTAATGACTACAGGAGCCTTCTGCAATTTGAAATTTTGGGTGCTCAT 262
DB 191 CA---GTATCATCATAGATCAGAGTAACGTAATTCACACTGCATTTTGGAGAATATTC 247
QY 263 GCCACAGGACAGAACAGAAATGAATAT-----TTGTTGACCTAGCTCTGAA 313
DB 248 GCTGTATGAATCGACACAGCTGAGAAATCGTTAAATGAGATTGCAATCTAGCTGTGAA 307
QY 314 TGCTCTCCAGGCTTATCAACCTGGCAATCAGTTATAAATAAATTAATTTTGTGTA 373
DB 308 TGCTATTCCAAACCTGTGCTGGGAGCAGCAAAACATTTCAAGACTTCTTTCTTCA 367
QY 374 AATAAGAGGAACCTTTAAATAATGATGTGTGAGAGCTTTTATACAATCAGACGCTTATGAA 433
DB 368 AGTAACGTGAGATTGAAAGATTTTGTAGGAGTATTGTACAACCAAGAAATTCATGGA 427
QY 434 GAGACTACAGAAACCAACTACGATTAATGCTTTATAGACCCCTGTGATCCCTGTGGAGA 493
DB 428 CAAGCTACGGATGACCAATATGATGATGAGTATAGACCCCTGTGCTGCTGCTGAGAGA 487
QY 494 CCTGATGGCTGAGTGTGCTTCAGTCCCTTTTGTGCTCACACTTAGAATTTCTGTAGGAG 553
DB 488 GTTGGTGGCAGAGTGTTCAGATCCCTTTTGTGATATACACTGAGGTTCAGCATGGGCTA 547
QY 554 CAATATGGAGGAGCTGTGGGAACCTTCCAGCTTCCACTTTTCCCTATGATACCTGTGCCTAT 613
DB 548 CTACATGGAGAAACACTGTGCCAGCTTCCAATTTCCACTCTCGTATGATACCGGTTGTCAT 607
QY 614 GACAGGACTACAGACAGATGACCTTTCTGGAAAGAGTAAAAATCAATGCTTTCAGT 673
DB 608 GAGTGAAGCTGACAGCAATATGACCTTCACAGAGAGGGTGAAGAAATATGATGTTTCACT 667
QY 674 TTTGTTCCACTTCTGGATTTCAGGATTCAGCTATCATTTTGGAGAGAGTTTATAGTAA 733
DB 668 GTTGTGTGACTACTGGCTCCAGCAATATGACTTTGCAATTCGGGATCAGTTTACAGTGA 727

Qy	734	GGCATTAGGAAGCCCACTACATTATGTGAGACTGTGGAAAAAGCTGAGATATGCGCTAAT	793	clone_lib:RIKEN full-length enriched mouse cDNA library clone:1300012D20.
Db	728	AACCCCTAGGAAGGCCCAACAGCTTCTGTAGACTGTGGGGAAGCTGACATTTGGCTAAT	787	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Qy	794	ACGAATATGGGATTTTGAATTTCTCAACCATACCAACCTAACTTTGAGTTTGTGG	853	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Db	788	CCGAATATTTGGGATTTTGGATTTCTCTGCTCCATATTTACCAAAATTTTGGTTTGGG	847	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Qy	854	AGCATTGCACTGTAAACCTGCCAAAGCTTTGGCTTAAGGAATGGAATAATTTTCTCCAGAG	913	Genome Res. 10 (10), 1617-1630 (2000)
Db	848	AGCATTGCACTGTAAACCTGCCAAAGCTTTTACCTTAAGGAATGGAATAATTTTCTCCAGAG	907	20499374 11042159
Qy	914	TTGAGGGAGAGATGGTATTTGTGTGTTTCTCTGGGCTCACTGTTTCAAAATTTTACAGA	973	REFERENCE
Db	908	CTCTGGAGACATGGTGTAGTACTATTTTCACTGGGTCAATGGTCAAAAACCTGCACAGA	967	REFERENCE
Qy	974	AGAAAAGGCTAATATATGCTTCAGCCCTTGGCCAGATCCCAAGAGGTGTATTGGAG	1033	REFERENCE
Db	968	AGAAAAGGCTAATATATGCTTCAGCCCTTGGCCAGATCCCAAGAGGTGTATTGGAG	1027	REFERENCE
Qy	1034	GTCAAAGGAAAAAACCATTAGGAGCAATACCTCGGCTGTATGATTGGATACC	1093	REFERENCE
Db	1028	ATACTAGGCAAGAGCCCAATATAGGATCCATCTCGGCTTTTAAATTTGGATTC	1087	REFERENCE
Qy	1094	CCAGAATGATCTTTTGGTCAATCCCAAAACCAAGCTTTTATCAGTCTGGTGAATGAA	1153	REFERENCE
Db	1088	CCAGAATGATCTTTTGGTCAATCCCAAAACCAAGCTTTTATCAGTCTGGTGAATGAA	1147	REFERENCE
Qy	1154	TGGATCTATGAAGCTATTTACCATGGGTCCTATGGTGGGAGTCCCATATTTGGTGA	1213	REFERENCE
Db	1148	CGGGATTTATGAAGCCATTTTACCATGGGTCCTATGGTGGGCTTCCCATGTTAGGGA	1207	REFERENCE
Qy	1214	TCAGCTTGATAACATAGCTCACATGAAGGCCAAGGAGCGCTAGTAATTAACATCA	1273	REFERENCE
Db	1208	TCAGCTTGATAACATAGCTCACATGAAGGCCAAGGAGCGCTAGTAATTAACATCA	1267	REFERENCE
Qy	1274	AACATGACAGGGAAGATTTACTGAGGGCTTTTGAACAAGTCAATACCCATTCCTCTTA	1333	REFERENCE
Db	1268	TACAATGACAGGACAGATTTACTGAGGGCTTTTGAACAAGTCAATACCCATTCCTCTTA	1327	REFERENCE
Qy	1334	TAAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTGTAAGGCCCTAGA	1393	REFERENCE
Db	1328	TAAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTGTAAGGCCCTAGA	1387	REFERENCE
Qy	1394	TCGAGCAGTCTTCTGGATCAGTTTGTGTCGCCCAAGGAGCAAGCACTCGCATC	1453	REFERENCE
Db	1388	CAGAGCAGTCTTCTGGATGAGTTTGTGTCGCCCAAGGAGCAAGCACTCGCATC	1447	REFERENCE
Qy	1454	AGTGCCCATGACCTCACCTGGTTCAGCAGCTACTCTATAGATGCTGATGGTTCCTGCT	1513	REFERENCE
Db	1448	GGCAGCCCATGACCTCACCTGGTTCAGTACCCTCCCTAGATGCTGATGGTTCCTATT	1507	REFERENCE
Qy	1514	GACCTGTGCGCACTGCTATATCTTGTTCACAAAATGTTTTTATTTCTGTCACAAA	1573	REFERENCE
Db	1508	GTGTGTGCTGCTGCTGATCTGATCTCATCATCAATAATTTTGTGTTGTGTGTCACAAA	1567	REFERENCE
Qy	1574	ATTATAAACTAGAAAGATAGAAAAGAGGGAATAGA	1611	REFERENCE
Db	1568	ACTTTATATGA-AGAAAAGTAAAGAAAATGGGGAACAGA	1604	REFERENCE
RESULT 2	AK004971	LOCUS	2573 bp mRNA linear	HTC 19-JAN-2002
DEFINITION	Mus musculus adult male liver cDNA, RIKEN full-length enriched			
KEYWORDS	2B1 PRECURSOR, MICROSMAL (EC 2.4.1.17) (UDPGT) (UDPGR-2), full			
SOURCE	Insert sequence.			
ACCESSION	AK004971			
VERSION	AK004971.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA,			

TITLE
JOURNAL

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of


```

Db      1522 TCCTATTGGTCTGTGTAGCTGTGGTATTTCATCATTTGCCAAATGTTGCCCTCTTTTCTT 1581
QY      1567 GTCAAAATTTAATAAACTAGAAAGATAGAAAA 1600
Db      1582 GCCATAAGACTGCTTAACATATGGGAAGAAGAAAA 1615

RESULT 3
LOCUS   BQ713091
DEFINITION AGENCOURT_8295152 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6309767
5', mRNA sequence.
ACCESSION BQ713091
VERSION   BQ713091.1 GI:21851990
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Susan L. Sullivan, PhD.
          cDNA Library Preparation: ResGen, Invitrogen Corp
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13730 row: p column: 24
          High quality sequence stop: 737.
          Location/Qualifiers
            1..960
              /organism="Mus musculus"
              /db_xref="taxon:10090"
              /clone="IMAGE:6309767"
              /clone_lib="NIH_MGC_129"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: olfactory epithelium; vector:
              PCMV-SPORT6.1.ccdB; Site_1: EcoRV; Site_2: NotI; Cloned
              unidirectionally. Primer: Oligo dT. Average insert size
              2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this
              is a NIH_MGC Library."
              273 a 204 c 237 g 246 t

BASE COUNT 273 a 204 c 237 g 246 t
ORIGIN

Query Match 38.2%; Score 625.2; DB 14; Length 960;
Best Local Similarity 83.0%; Pred. No. 3.le-141;
Matches 725; Conservative 0; Mismatches 148; Indels. 1; Gaps 1;

QY      740 AGGAAGCCCACTACATATTATGTAGACTGTGGAAAGCTGAGATATGCTTAATACGAAC 799
Db      45 AGGAAGCCCACTACATATTATGTAGACTGTGGAAAGCTGAGATATGCTTAATACGAAC 104
QY      800 ATATTGGGATTTGAATTTCTCAACCACTTACTTTTGAGTTTGTGGAGGATT 859
Db      105 CTATTGGGATTTGAATTTCTCGTCTATATTACCGAATTTTGAGTTTGTGGAGGACT 164
QY      860 GCACTGTAACCTGCCAAAGCTTTGCGCTAAGAAATGGAATTTTGTCCAGAGTTCCAGG 919
Db      165 GCACCTCAACCTGCCAAGCCTTTACCTAAGAAATGGAAGATTTGTCCAGACCTCAGG 224
QY      920 GGAAGATGTTATTTGGTGTCTCTCTGGGGTCACTGTTTCAAAATGTTACAGAGAAAA 979
Db      225 GGAACATGTTATTTGGTGTCTCTCTGGGGTCAATGGTCAAAACCTGCAGCATGAAAA 284
QY      980 GCCTAATATCATTTGCTCAGCCCTTGCAGATCCCAAGAGGTGTTATGAGGTACAA 1039
Db      285 GCCCAATCTATTTGCTCAGCCCTTGCAGATCCCTCAGAGATTCCTCAGAGGTTTGTGGCGATACA 344

1040 AGAAAAAACCACATCCACATTAGGAGCAATACTCGGCTGTATGATTGGATACCCAGAA 1099
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
345 AGGAAGATACAGACACATTAGATCCAATACAAGACTGTTGATGGATTCCTCAGAA 404
1100 TGATCTTTTGGTCATCCCAAAACCAAGCTTTTATCACTCATGTTGGATGAATGGGAT 1159
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
405 TGATCTTTTGGACATCCCAAAACCAAGAGCTTTTATCACCCATGTTGGAAACAAATGAAT 464
1160 CTATGAAGCTATTTACCATGGGCTCTATGTTGGAGTTCCTCATATTTGGTGATCAGCT 1219
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
465 CTATGAGGCTATTTACCATGGGATCCCTATGTTGGAGTGCCCATGTTCCGTGACCAAGC 524
1220 TGATAACATAGCTACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAAACTAT 1279
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
525 TGACAACATTTGCTACATGAAGGCCAAAGGAGCAGCAGTGGAGTGAACATGACACGAT 584
1280 GACAAAGGAGATTTTACTGAGGCTTTGAGACACAGTCATACCGATTCCTCTTATAAAGA 1339
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
585 GACAAGCTCAGATCTGCTCAATGCTGTGAGAACAGTCATCAATGAGCCATCTTATAAAGA 644
1340 GAATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGC 1399
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
645 GAATGCCATGAGACTATCAAGATCCACCATGACCAGCCAGTGAAGCCCTCGACCGAGC 704
1400 AGTCTTCTGGATCGAGTTTGTGTCATGGCCACAAAGGAGCAGCAACCTGCGATCAGCTGC 1459
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
705 AGTCTTCTGGATCGAGTTTGTGTCATGGCCACAAAGGAGCAGCAACCTGCGTGTGCGAGC 764
1460 CCATGACCTCAGCTGTTCCACACTACTCTATAGATGTGATGGTTCCTGCTGCTGACCTG 1519
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
765 CCATGACCTCAGCTGTTCCACACTACTCTGTCAGCTCTCTGATGTAATGGGTTTCTGCTGGGCTG 824
1520 TGTGCAACTGCTATATTCTTGTTCACAAATGTTTTTATTTT-TCCTGTCAAAATTTTA 1578
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
825 TGTGCACTGCTATATTATTGTTGCAAAATGCTGTTGTTTATATAATTTCAAAAGTTG 884
1579 ATAAACTAGAAAGATAGAAAGAGGGAATAGAT 1612
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
885 GTAAGACAGGAAGAGAGAAAAAACGTGACTAGTT 918

RESULT 4
LOCUS   BQ942104
DEFINITION AGENCOURT_87631113 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6313348
5', mRNA sequence.
ACCESSION BQ942104
VERSION   BQ942104.1 GI:22357582
KEYWORDS EST.
SOURCE   house mouse.
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Susan L. Sullivan, PhD.
          cDNA Library Preparation: ResGen, Invitrogen Corp
          DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13740 row: f column: 05
          High quality sequence stop: 656.
          Location/Qualifiers
            1..1004
              /organism="Mus musculus"
              /db_xref="taxon:10090"
              /clone="IMAGE:6313348"

FEATURES
          source

```



```
/clone_lib="NH_MGC_129"
/lab_host="DH10B (phage-resistant)"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site_1: EcoRV; Site_2: NotI; Cloned
unidirectionally. Primer: Oligo dr. Average insert size
2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this
is a NH_MGC Library."
BASE COUNT      286 a      219 c      249 g      249 t      1 others
ORIGIN

Query Match      38.1%; Score 623.2; DB 14; Length 1004;
Best Local Similarity 81.9%; Pred. NO. 9.5e-141;
Matches 730; Conservative 0; Mismatches 159; Indels 2; Gaps 1;

QY 714 TGGGAAGAGTTTATAGTAAGGCATATAGGAAGGCCACTACATATTGTGAGACTGTGGGA 773
DB 714 TGGGAAGAGTTTATAGTAAGGCATATAGGAAGGCCACTACATATTGTGAGACTGTGGGA 773
QY 774 AAAGCTGAGATATGGCTTAATACGACATATTTGGGATTTTGAATTTCTCAACCATACCA 833
DB 774 AAAGCTGAGATATGGCTTAATACGACATATTTGGGATTTTGAATTTCTCAACCATACCA 833
QY 113 AAAGCTGAGATTTGGCTAATCGAACCTATTGGGATTTTGAATTTCTCGTCCATATTTA 172
DB 113 AAAGCTGAGATTTGGCTAATCGAACCTATTGGGATTTTGAATTTCTCGTCCATATTTA 172
QY 834 CCTAACCTTGGATTTGTTGGAGGATTCACCTGTAACTGCAAGCTTTGCCCTAAGGA 893
DB 834 CCTAACCTTGGATTTGTTGGAGGATTCACCTGTAACTGCAAGCTTTGCCCTAAGGA 893
QY 173 CCGAATTTTGATTTGTTGGGAGGACTGCACCTGCAAACTGCGCAAGCTTTTACCTAAGGA 232
DB 173 CCGAATTTTGATTTGTTGGGAGGACTGCACCTGCAAACTGCGCAAGCTTTTACCTAAGGA 232
QY 894 ATGGAATTTTGTCCAGATTCAGGGAAGATGTTATGTTGGTGTCTTCTCTGGGTCA 953
DB 894 ATGGAATTTTGTCCAGATTCAGGGAAGATGTTATGTTGGTGTCTTCTCTGGGTCA 953
QY 233 ATGGAAGATTTTGTCCAGACCTCAGGGAACATGTTATGTTGGTGTCTTCTCTGGGTCA 292
DB 233 ATGGAAGATTTTGTCCAGACCTCAGGGAACATGTTATGTTGGTGTCTTCTCTGGGTCA 292
QY 954 CTGTTTCAAAATGTTACAGAGAAGAGCTAATATCATGTTGTTGAGCCCTTGCACAGATC 1013
DB 954 CTGTTTCAAAATGTTACAGAGAAGAGCTAATATCATGTTGTTGAGCCCTTGCACAGATC 1013
QY 293 ATGTTCAAAACCTCAGACATGAAAGGCCAATCTCATTTGCTCAGCCCTTGCACAGAT 352
DB 293 ATGTTCAAAACCTCAGACATGAAAGGCCAATCTCATTTGCTCAGCCCTTGCACAGAT 352
QY 1014 CCAGAGAGGTTATGAGGTACAGAGGAAAGAAACCATCCATATAGGAGCAATACT 1073
DB 1014 CCAGAGAGGTTATGAGGTACAGAGGAAAGAAACCATCCATATAGGAGCAATACT 1073
QY 353 CCTCAGAGGTTTGTGCGGATACAGAGGAAAGATACACACATATAGATTCACATACA 412
DB 353 CCTCAGAGGTTTGTGCGGATACAGAGGAAAGATACACACATATAGATTCACATACA 412
QY 1074 CGGCTGTATGATGGATACCCAGATGATCTCTTGTGTCATCCAAACCAAGAGCTTTT 1133
DB 1074 CGGCTGTATGATGGATACCCAGATGATCTCTTGTGTCATCCAAACCAAGAGCTTTT 1133
QY 413 AGACTGTTTGGATTTGCTCAGATGATCTCTTGTGACATCCCAAAACCCAGAGCTTT 472
DB 413 AGACTGTTTGGATTTGCTCAGATGATCTCTTGTGACATCCCAAAACCCAGAGCTTT 472
QY 1134 ATCACTCATGTGGAATGGAATGGAATCTATGAAGCTATTTACCATGGGTCCCTATGGTG 1193
DB 1134 ATCACTCATGTGGAATGGAATGGAATCTATGAAGCTATTTACCATGGGTCCCTATGGTG 1193
QY 473 ATCACTCATGTGGAATGGAATGGAATCTATGAAGCTATTTACCATGGGTCCCTATGGTG 532
DB 473 ATCACTCATGTGGAATGGAATGGAATCTATGAAGCTATTTACCATGGGTCCCTATGGTG 532
QY 1194 GGAGTTCCTCATTTTGGTGTGATGCTGATTAACATAGCTACATGAAGGCCAAAGGAGCA 1253
DB 1194 GGAGTTCCTCATTTTGGTGTGATGCTGATTAACATAGCTACATGAAGGCCAAAGGAGCA 1253
QY 533 GGAGTGCCTCATTTGCTGACCGCTGACAAACATGCTCATGAAGGCCAAAGGAGCA 592
DB 533 GGAGTGCCTCATTTGCTGACCGCTGACAAACATGCTCATGAAGGCCAAAGGAGCA 592
QY 1254 GCTGTAGAAATTAACCTTCAAACTATGACAGCGAAGATTTTACTGAGGCTTTGAGAAC 1313
DB 1254 GCTGTAGAAATTAACCTTCAAACTATGACAGCGAAGATTTTACTGAGGCTTTGAGAAC 1313
QY 593 GCAGTGGAGGTGAACATGAACACGATGACAGCTCAGATGCTGCTCAATGCTGAGAAC 652
DB 593 GCAGTGGAGGTGAACATGAACACGATGACAGCTCAGATGCTGCTCAATGCTGAGAAC 652
QY 1314 GTCATTACCGATTCCTTTAAGAGAGATGCTATGAGATTTATCAAGATTCACCATGAT 1373
DB 1314 GTCATTACCGATTCCTTTAAGAGAGATGCTATGAGATTTATCAAGATTCACCATGAT 1373
QY 653 GTCATCAATGAGCCATCTTTAAGAGAGATGCTATGAGATTTATCAAGATTCACCATGAT 712
DB 653 GTCATCAATGAGCCATCTTTAAGAGAGATGCTATGAGATTTATCAAGATTCACCATGAT 712
QY 1374 CAACCTGTAAAGCCCTTATGATGAGCAGCTCTTCTGGATGCGAGTTGTTGTCATGCGCACAAA 1433
DB 1374 CAACCTGTAAAGCCCTTATGATGAGCAGCTCTTCTGGATGCGAGTTGTTGTCATGCGCACAAA 1433
QY 713 CAGCCAGTGAAGCCCTTGGACCGGAGCAGCTCTTCTGGATGAGTTGTCATGCGTCACAAA 772
DB 713 CAGCCAGTGAAGCCCTTGGACCGGAGCAGCTCTTCTGGATGAGTTGTCATGCGTCACAAA 772
QY 1434 GGAGCCAGACCTCGGATACAGCTGCCCCATGACCTACCTCGTTTCCAGCAGCTACTCTATA 1493
DB 1434 GGAGCCAGACCTCGGATACAGCTGCCCCATGACCTACCTCGTTTCCAGCAGCTACTCTATA 1493
QY 773 GGAGCCAGACCTCTCGTGTGGCAGCCCATGACCTCAGCTGTTTCACTACCATCTCTG 832
DB 773 GGAGCCAGACCTCTCGTGTGGCAGCCCATGACCTCAGCTGTTTCACTACCATCTCTG 832
QY 1494 GATGTGATTTGGTTCCTGCTGACCTGTGTGCACTGCTATATTCCTT--GTTCACAAAAT 1551
DB 1494 GATGTGATTTGGTTCCTGCTGACCTGTGTGCACTGCTATATTCCTT--GTTCACAAAAT 1551
QY 833 AATGGAATTTGGGTTCTGCTGGCCTGTGTGGCATCTGCTAAATATTGGGTGCCAAAAGG 892
DB 833 AATGGAATTTGGGTTCTGCTGGCCTGTGTGGCATCTGCTAAATATTGGGTGCCAAAAGG 892
QY 1552 GTTTTTTTTCTGTCACAAAATTTTATAAACTAGAAAAGATAGAAAAGA 1602
DB 1552 GTTTTTTTTCTGTCACAAAATTTTATAAACTAGAAAAGATAGAAAAGA 1602
QY 893 GTGGTGTGTTTAAATTCAAAAGAGTTGTTAGAACCCGGAAGAAAAGAAA 943
DB 893 GTGGTGTGTTTAAATTCAAAAGAGTTGTTAGAACCCGGAAGAAAAGAAA 943
```

RESULT 5
AK002736
LOCUS

DEFINITION

AK002736 1896 bp mRNA linear HTC 19-JAN-2002
Mus musculus adult male kidney cDNA, RIKEN full-length enriched
library, clone:0610033E06:UDP-glucuronosyltransferase 2 family,
member 5, full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE


```

||||| 1138 AAACCAAGCCCTTTATAACTCATGTGGGGCCCAACAGTCTCTATGAGGCAATCTATCATG 1197
QY ||||||| 1180 GGGTCCCTATGTTGGGAGTCCCATATATTGGTGATCAGCTTGATGAACATAGCTCACATGA 1239
Db ||||||| 1198 GAATCCCTATGATGGCATTCCTTTGTTGGGAACAGCATGATAACATGCCCCATATGG 1257
QY ||||||| 1240 AGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAACTATGACAAAGCGAAGATTCTACTGA 1299
Db ||||||| 1258 TGGCCAAAGGAGCAGCTGTATCATTTGAATATCAGAACATGTCAAGGTGAGATGCTCA 1317
QY ||||||| 1300 GGGCTTTGAGAACACATCAATACCGATTCTCTTATAAGAGAAATGCTATGAGATPATCAA 1359
Db ||||||| 1318 ATGCACTGGAGGAAGTCAAGAAATCCTTTCTATAAAAGAAATGCTATGTTGTTGTCAA 1377
QY ||||||| 1360 GAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGTCTCTGGATCGAGTTTG 1419
Db ||||||| 1378 CCATTACCATGACGAGCTATGAACCCCTGGACAGGCTGTATCTGGGTGAGTTTG 1437
QY ||||||| 1420 TCATCGCCACAAAGGAGCCAGCACCTCGCATCAGCTGCCCATGACCTCACTCGTGTTC 1479
Db ||||||| 1438 TCATCGCCACAAAGTCTTAAGCAGCTTGAAGCCACTTGGACATAACCTTACCTGGTACC 1497
QY ||||||| 1480 AGCACTACTCTATAGATGTGATGGTTCCTGCTGACCTGTGTGGCAACTGCTATATCT 1539
Db ||||||| 1498 AGTACCACTCTCTGGATGTGATGGATTCTTACTCGCTTGTGTGGCAGCCACTGTGATTC 1557
QY ||||||| 1540 TCTTCACAAATGTTTTTATTTTCTCTGTCACAAATTTT---AATAAAACTAGAAAGATAG 1596
Db ||||||| 1558 TTACTGTAAAGTCCCTCTGTGTCATTTACCGATTCTTTGTAACGAAAAAAGAAATG 1617
QY ||||||| 1597 AAAGAGGGAATAGATCTTTTCCAAATTCAGAAA 1630
Db ||||||| 1618 AAGAATGAGTAGATGTCATTTGTCATATGTCACACACA 1651

```

RESULT 6

BQ925596
LOCUS
DEFINITION AGENCOURT_8888479 NIH_MGC_129 Mus musculus cdna clone image:6311424
5', mRNA sequence.

ACCESSION BQ925596
VERSION BQ925596.1 GI:22340627
KEYWORDS EST.

SOURCE

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 927)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM13735 row: f column: 01
High quality sequence stop: 635.
Location/Qualifiers

FEATURES

source

1..927
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6311424"
/clone_lib="NIH_MGC_129"
/lab_host="DH10B (phage-resistant)"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1 ccdb; Site:1; EcorV; Site:2; NotI; Cloned
unidirectionally. Primer: Oligo dt. Average insert size

2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this
is a NIH_MGC Library."
BASE COUNT 262 a 201 c 230 g 231 t 3 others
ORIGIN

Query Match 36.7%; Score 600; DB 14; Length 927;
Best Local Similarity 80.4%; Pred. No. 4.1e-135;
Matches 714; Conservative 0; Mismatches 172; Indels 2; Gaps 1;

```

QY 740 AGGAAGGCCCACTACATATGTGAGACTGTGGGAAAGCTGAGATATGCTTAATACGAC 799
Db ||||||| 37 AGGAAGGCCCACTACATATGTGAGACTGTGGGAAAGCTGAGATATGCTTAATACGAC 96
QY 800 ATATTGGGATTTGAATTTCTCAACCATACCACTTAACCTTTGAGTTGTTGGAGGATT 859
Db ||||||| 97 CTATTGGGATTTGAATTTCTCGTCCATATTACCGAATTTTGAATTTGGGAGGACT 156
QY 860 GCATGTAAACCTGCCAAGCTTTGCCCTAAGAAATGGAATTTTGTCCAGAGTTACGG 919
Db ||||||| 157 GCATGTCAACCTGCCAAGCTTTTACCTAAGAAATGGAAGAGTTTGTCCAGACCTCAGG 216
QY 920 GGAAGATGATTTGGTGTGTTTCTCTGGGTCACTGTTTCAAAATGTTACAGAAAGAAA 979
Db ||||||| 217 GGAACATGGTATTTGGTGTGTTTCTCTGGGTCAATGTTCAAAACCTTGACAGATGA 276
QY 980 GGCTAATATCATTTGCTTCAGCCCTTTGCCAGATCCCAAGAGGTGTTTATGGAGTACAA 1039
Db ||||||| 277 GGCCAAATCTATTTGCTTCAGCCCTTTGCCAGATTTCTCAGAAGGTTTTTGTGGCGATA 336
QY 1040 AGGAAGAACCATCCATAGGAGCCCAATCTCGGCTGTATGATTTGGATACCCAGAA 1099
Db ||||||| 337 AGGAAGATACCAACATAGGATTTAGGATTTCAACAGAGTGTGTTGATTTGGATTTCTCA 396
QY 1100 TGATCTTCTTGCTATCCCAAAACCAAGCTTTTATCACTCATGTGTGAATGAATGGAT 1159
Db ||||||| 397 TGATCTTCTTGACATCCCAAAACCAAGCTTTTATCACTCATGTGTGAATGAATGGAT 456
QY 1160 CTATGAAGCTATTTACCATTGGGTCCCTATGTTGGAGTTCCTCATATTTTGGTGTGAT 1219
Db ||||||| 457 CTATGAGGCTATTTACCATTGGGTCCCTATGTTGGAGTTCCTCATATTTTGGTGTGAT 516
QY 1220 TGATAACATAGCTCATCAAGGCCCAAGGAGGAGCAGCTGTAGAAATTAATCTTCAAACTAT 1279
Db ||||||| 517 TGACAAATTTGCTACATGAAGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 576
QY 1280 GACAAAGCAAGATTTACTGAGGCTTTGAGAACACTCATTTACCGATTTCTCTTATAAGA 1339
Db ||||||| 577 GACAAAGCTAGATCTGCTCAATGCTCTGAGAACACTCATCATGAGCCATCTTATNAAGA 636
QY 1340 GAATGCTATGAGATTATCAAGAAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGC 1399
Db ||||||| 637 GAATGCCATGAGACTATCAAGAAATCCACCATGACCAAGCAGTGAAGCCCTCGGAGCCGAGC 696
QY 1400 AGTCTTCTGGATCGAGTTTTCATCGCCACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1459
Db ||||||| 697 AGTCTTCTGGATTTGAGTTTGTCTGACANAGAGGCCAGCAGCAGCAGCAGCAGCAGCAGC 756
QY 1460 CCATGACCTCAGCTGTTTCCAGCAGTACTCTATAGATGTGATTTGGTGTCTGCTGACCTG 1519
Db ||||||| 757 CCATGACCTCAGCTGTTTTCAGTACCACTCTCCGGATGGAATTTGTTTCTGCTGCTGCTG 816
QY 1520 TGTGGCAACTGCTATTTCTTT--GTTTCAAAATTTTATTTTATTTTCTGCTGCAAAATTT 1577
Db ||||||| 817 TGTGGCATCTGCTATAATTTTGGGTGTCAAAGGAGTGTGTTTTCGTTTAAATTTTCCAAAA 876
QY 1578 AATAAACTAGAAAGATAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1625
Db ||||||| 877 AGTTGGGTAGGACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924

```

RESULT 7

BQ934372

LOCUS

BQ934372 930 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT_8764945 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6314635
5', mRNA sequence.
ACCESSION B0934372
VERSION B0934372.1 GI:22349755
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 930)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Resgen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13743 row: k column: 20
High quality sequence stop: 547.
Location/Qualifiers
1. .930
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6314635"
/lab_host="NIH_MGC_129"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site.1: EcorV; Site.2: NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this
is a NIH_MGC Library."
BASE COUNT 275 a 200 c 230 g 225 t
ORIGIN
Query Match 33.8%; Score 552.4; DB 14; Length 930;
Best Local Similarity 81.5%; Pred. No. 1.5e-123;
Matches 665; Conservative 0; Mismatches 146; Indels 5; Gaps 2;
Qy 717 GAAGACTTTTATAGTAAAGCATAGGAGCCCACTACATATGTGAGACTGTGGGAAA 776
Db 3 GGACACTTGGGTTTAAAGTACAGGAGACCCACACACTGTGTGAGACTATGGGAAA 62
Qy 777 GCTGAGATATGCTATACGAACATATGGGATTTTGAATTTCTCAACCATACCAACT 836
Db 63 GCTGAGATTTGGCTAATGCAACCTATTGGGATTTTGAATTTCTCGTCCATATTTACCG 122
Qy 837 AACTTTGAGTTTCTGGAGATTCACCTGTAAACCTGCCAAAGCTTTGCTTAAGGAATG 896
Db 123 AATTTTGGTTTGGGGAGCTGCACTGCAAAACCTGCCAAAGCTTTTACCTAAGGAATG 182
Qy 897 GAAATTTTCTCAGAGTTTCAGGGAGATGGTATGTGGTGTCTCTGGGTCAGTG 956
Db 183 GAAGAGTTTCTCAGACCTCAGGGAGACATGGTATGTGGTGTCTCTGGGTCAGTG 242
Qy 957 TTTCAAAATTTACAGAGAAAAGGTATATATCATTTTCAGCCCTTGGCCAGATCCCA 1016
Db 243 GTCAAAACCTGCAGATGAAAAGGCAATCTCATTCGCTCAGCCCTTGGCCAGATTCCT 302
Qy 1017 CAGAGGTGTATGGAGTACAAGGAAAAAACCATCCACATAGGAGCCATACCTCGG 1076
Db 303 CAGAAGTTTGTGGGATACAAGGAAGATACAGACACATTAGATCCATACAAAGA 362
Qy 1077 CTGTATGATGGATACCCAGATGATCTTCTTGGTCATCCCAAAACCAAGCTTTTATC 1136
Db 363 CTGTTGATTTGGATTCTCAGATGATCTTCTTGGATCCCAAAACCAAGCTTTTATC 422
Qy 1137 ACTCATGGTGAATGAATGGATCTATGAAGCTATTTACCATGGGTCCCTATGGTGGGA 1196
Db 1137 ACTCATGGTGAATGAATGGATCTATGAAGCTATTTACCATGGGTCCCTATGGTGGGA 1196

423 ACCATGGTGGACAAATGGAATCTATGAGGCTATTATACCATGGGATCCCTATGGTGGGA 482
1197 GTTCCCATATTTGGTGATCAGCTTGATACATAGCTCAGATGAAGCCAAAGCAGCAGCT 1256
483 GTGCCCATGTTCGCTGACCAGCCTGACAACATTTGCTCAGATGAAGCCAAAGGAGCAGCA 542
1257 GTAGAAATAAATCTTCAAAACTATGACAAGCAGAGATTTACTGAGGGCTTTTGAACAAGTC 1316
543 GTGGAGGTGAACATGAACAGATGACAAGCTCAGATCTGCTCAATGCTCTGAGAACAGTC 602
1317 ATTACCGATTCTCTTTATAAAGAGATGCTATGAGATTTATCAAGAATTCACCATATCAAA 1376
603 ATCAATGAGCCATCTTTATAAAGAGATGCGATGAGACTATCAAGAATTCACCATGACCAG 662
1377 CCTGTAAAGCCCTAGATGAGCAGTCTTCTGTGATGAGTTTCTCATGCGCCCAAGAAGGA 1436
663 CCAGTGAAGCCCTGAGCAGCAGTCTTCTGTGATGAGTTTCTCATGCGTCAACAAGGA 722
1437 GCCAAGCAGCTGGGATCAGCTGCCAT-GACCTACCTGCTTCC-----AGCAGTACTCTA 1491
723 GCCAAGCAGCTTGGTGTGGCAGCCCATGCCCTCAGCTGGTTTCCAGAACCCCTCTCTG 782
1492 TAGATGTGATGGGTTCTGCTGACCTGTGTGGCAA 1527
783 GATGTAATGGGGTTCTCTGCTGCTGCGCTGTGGGCCA 818

RESULT 8
BQ714919
LOCUS
DEFINITION BQ714919 1030 bp mRNA linear EST 16-JUL-2002
AGENCOURT_8480154 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310837
5', mRNA sequence.
ACCESSION BQ714919
VERSION BQ714919.1 GI:21853818
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1030)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Resgen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13733 row: m column: 14
High quality sequence stop: 599.
Location/Qualifiers
1. .1030
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6310837"
/lab_host="NIH_MGC_129"
/note="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site.1: EcorV; Site.2: NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this
is a NIH_MGC Library."
BASE COUNT 295 a 229 c 257 g 248 t
ORIGIN
Query Match 31.8%; Score 520.4; DB 14; Length 1030;
Best Local Similarity 78.5%; Pred. No. 8.8e-116;
Matches 724; Conservative 0; Mismatches 186; Indels 12; Gaps 8;

```
Qy 717 GAAGAGTTTTATAGTAGGCAATAGGAAGGCCCACTACATATATCTGAGACTGTGGGAAA 776
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 777 GCTCAGATATGGCTAATACGACATATATGGATTTTGAATTTCTCCTCAACCATACCACT 836
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 86  GCTCAGATATGGCTAATACGACATATATGGATTTTGAATTTCTCCTCCTCAATATTTACCG 145
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 837 AACTTTCAGTTTGTGGAGATTCACCTGTAAACCTGCCAAGCTTGGCTTAAGGAATG 896
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 146 AATTTTCAGTTTGTGGAGATTCACCTGTAAACCTGCCAAGCTTGGCTTAAGGAATG 205
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 897 GAAATTTTTCAGATTCAGGGAAGATGATTTGTGGTGTCTCTGGGGTCACTG 956
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 206 GAAGAGTTTGTCCAGACCTCAGGGAACATGATTTGTGGTGTCTCTGGGGTCAATG 265
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 957 TTTCAAATCTTACAGAGAAAGGCTAATATCATCTTCTCAGCCCTTGCCCATCCCA 1016
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 266 GTCAAAACCTTGACAGATGAAGGCCAATCTCATTTGCCCTCAGCCCTTGCCCATCCCT 325
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1017 CAGAAGTGTATGGAGGTACAAAAGGAAACCAACCATCCACATTTAGGAGCCAATACCTCGG 1076
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 326 CAGAAGTGTATGGAGGTACAAAAGGAAACCAACCATCCACATTTAGGAGCCAATACCAAGA 385
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1077 CTGATGATGGATACCCCAAGATGATCTTCTGGTCATCCCAACCAAGACCTTTATC 1136
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 386 CTGTTGATGATCTCTCAGATGATCTCTTGGACATCCCAACCAAGACGCTTTATC 445
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1137 ACTCATGGTGAATGATGATCTATGAGCTATTTACCATGGGTCCCTATGGTGGGA 1196
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 446 ACCATGGTGAACAAATGGAATCTATGAGGCTATTTACCATGGGTCCCTATGGTGGGA 505
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1197 GTTCCCATATTTGTGTGATGATGATAAGCTACATGATGATGATGATGATGATGATGATGAT 1256
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 506 GTGCCCATGTCGTGACCGCTGACACATCTCTCAGATGAGGCAAGGAGGACGA 565
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1257 GTAGAAATAAATCTCAAACTATGACAAAGCAAGATTTACTGAGGGCTTTGAGAACAGTC 1316
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 566 GTGAGGTGAACATGAACAGCATGACAAAGCTCAGATCTGCTCAATGCTCTGAGAACAGTC 625
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1317 ATTACCATGCC-TCTTATGAAGAGATGCTATGAGATTTATCAGATTTACCATGATCA 1375
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 626 ATCAATGAGCCCATCTTATGAAGAGATGCGCATGAGACTATCAAGAATCCCCATGACCA 685
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1376 ACCTGTAAAGCCCTTAGA-TCGAGCAGTCTTCTGGATCGAG-TTGTGTCATCGCCCAAAA 1433
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 686 GCCAGTGAAGCCCTTGACCCGAGCAGTCTCTGATGATGATGATGATGATGATGATGATGAT 745
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1434 GG-AGCCAGCAGCTGCGATCAGCTGCCATGACCTCACCT--GGTCCAGCAGTACTCT 1490
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 746 GGAAGCCAGCAGCCTTCGGGTGGCAGCCCATGACCTCAGCTTGGGTTCAGAGCCCACTCT 805
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1491 ATAGATGTA--TTGGGTTCCTGCTGACCTGTGTGGCAA-CTGCTATATTTCTCTT--C 1544
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 806 CTGATGGAATTTGGGTTCCTGCTGACCTGTGTGGCAA-CTGCTATATTTCTCTT--C 865
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1545 ACAAAATGTTTTTTTATTTCTGCAAAAATTTTAAATAAAGTAGAAGATAGAAAGAGG 1604
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 866 AAAAGGTGGTGTGTTTATTTCTTATATTTCTCCCAAGTTGGTAACCCGGAAGAGAAAAAACT 925
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1605 GAATAGATCTTTCCAAATCAA 1626
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 926 GGACTAATTTGAAGAAAAAAA 947
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
BQ714368 913 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8468705 NIH_MGC_129 Mus musculus cdna clone IMAGE:6308241
DEFINITION 5', mRNA sequence.
ACCESSION BQ714368
VERSION BQ714368.1 GI:21853267
KEYWORDS EST.
```

```
house mouse.
Mus musculus
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 913)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, Ph.D.
cDNA Library Preparation: Resgen, Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW13727 row: a column: 10
High quality sequence stop: 654.
Location/Qualifiers
1. 913
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6308241"
/lab_host="NIH_MGC_129"
/notes="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site:1; EcorV; Site:2; NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this
is a NIH_MGC Library."
BASE COUNT 258 a 221 c 203 g 231 t
ORIGIN
Query Match 30.1%; Score 492.6; DB 14; Length 913;
Best Local Similarity 81.0%; Pred. No. 5e-109;
Matches 573; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
Qy 919 GGGAGAGATGGTATTTCTCTCTGGGTCTGTTTCTCAAAATTTACAGAGAAA 978
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 979 AGGCTAATATCATTTGCTTACGCCCTTGCCAGATCCCAAGAGGTGTTTATGGAGTACA 1038
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 64  AGGCCAATCTCATTCGCTCAGCCCTTGCCAGATTTCTCAGAAGGTTTGTGGGATACA 123
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1039 AAGGAAAAACCATCCACATTTAGGAGCAATCTCGCTCTGATGATGATGATGATGATGATGATGAT 1098
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 124  AAGGAAAGATACCAAGACATTTAGGATCCAATACAAGACTGTTTGTATGGATTTCTCAGA 183
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1099 ATGATCTTCTTGGTCTATCCCAACCAAGCTTTTATCCTCATCTGTTGGTGAATGAATGGGA 1158
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184  ATGATCTTCTTGGATCCCAACCAAGAGCTTTTATCACCCTATGTTGGTGAACAAATGAA 243
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1159 TCTATGAAGCTATTTACCATGGGTCCTGATGGTGGAGTTCCTCATATTTGCTGATCAGC 1218
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244  TCTATGAGGCTATTTACCATGGGATCCCTATGTTGGGAGTCCCATGTTCTGCTGACCA 303
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1219 TTGATAACATAGCTCAGATGAAGCCAAAGGAGCAGCTGTAGAAATAAATTAACCTTCAAACTA 1278
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304  CTGACACATATGCTCAGATGAAGCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 363
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1279 TGACAAGCGAAGATTTTACTGAGGCTTTTGAGAAGCAGTCAATTTACCGATTTCTCTTTATAAG 1338
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 364  TGACAAGCTCAGATCTGCTCAATGCTCTGAGAAGCAGTCAATGAGGAGGAGGAGGAGGAGGAGGAG 423
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1339 AGAATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAG 1398
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 424  AGAATGCCATGAGATCTATCAAGAATTCACCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1399 CAGTCTTCTGATCGAGTTTGTCTATGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1458
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 484  CAGTCTTCTGATGAGTTTGTCTATGCTCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 543
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 1459 CCCATGACCTCACCTGGTCCAGCAGCTACTCTATAGATGATGGGTTCTGCTGACCT 1518
 DB 544 CCCATGACCTCAGCTGGTTTCAGTACCCTCTCTGGATGTAATGGGTTCTGCTGACCT 603
 QY 1519 GTGTGGCACTGCTATATCTTGTTCACAAATGTTTTTATTTCTCTGTCACAAATTTA 1578
 DB 604 GTGTGGCATCTGCTATATATTTGTTTCCAAATGTTGTTTATTTATTTCAAAAGTTG 663
 QY 1579 ATAAACTAGAAAGATAGAAAGAGGGAATAGATCTTTTCCAAATTTCA 1625
 DB 664 GTAAGACAGGAAGAAGAAAACGCTGACTAGTTGAAGACAAAGACA 710

RESULT 10
 BF689099/c
 LOCUS
 DEFINITION 6021851721 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4299880 3',
 mRNA sequence.
 ACCESSION BF689099
 VERSION BF689099.1 GI:11974507
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHC1157 row: 0 column: 17
 High quality sequence start: 5
 High quality sequence stop: 796.
 Location/Qualifiers
 1..823
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4299880"
 /clone_lib="NIH_MGC_43"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library. |"
 221 a 199 c 164 g 239 t

BASE COUNT
 ORIGIN
 Query Match 29.2%; Score 477; DB 12: Length 823;
 Best Local Similarity 78.0%; Pred. No. 3.1e-105;
 Matches 587; Conservative 0; Mismatches 165; Indels 1; Gaps 1;

QY 865 GTAAACCTGCCAAAGCTTTGCTTAGGAAATGGAAATTTTGTCCAGAGTTCCAGGGAAG 924
 DB 804 GCAAACCGCCCAACCCCTGCCTAGGGAATGGAGACTTTGTACAGATTCTCGAGAAA 745
 QY 925 ATGGTATTGTGCTGTTTCTCTGGGGTCAGTGTTCATAAAATGTTACAGAGAAAGGCTA 984
 DB 744 ATGGGTATTGTGCTGTTTCTCTGGGGTCAATGGTCAGTAACATGACAGAGAAAGGCCA 685

QY 985 ATATCATTTGCTTCAGCCCTTGGCCAGATCCACAGAGCTGTATGAGGTACAAAGGNA 1044
 DB 684 AGGTATTGCTATCAGACCTGGCCAGATCCCAAAAGGTTCTGTGGAGATTGATGGGA 625
 QY 1045 AAAACCATCCACATTTAGGAGGCAATPACTCGCTGTATGATGGATACCCCAAGATGATC 1104
 DB 624 ATAAACCATGATACCTTAGGTCTCAATACTCGCTGTATAAGTTGGATACCCCAAGATGAC 565
 QY 1105 TTCTTGGTCATCCCAAAACCAAGCTTTTATCATTCTGATGCTGGTGGATGAATGGATCTATG 1164
 DB 564 TTCTAGGTCTATCCAAAGACAGAGCTTTTATAACTCATGTTGGTGGAGCCCAATGGCATCTACG 505
 QY 1165 AAGCTATTACCATGGGTCCCTATGTTGGGAGTTTCCCATATTTTGGTGTGATCAGCTTGATA 1224
 DB 504 AGGCAATCTACCATGGGATCCCTATGTTGGGAGTTTCCATTTGTTCCCGATCAACCTGATA 445
 QY 1225 ACATAGCTCACATGAAGCCCAAGAGGAGCAGCTGTAGATAATAACTTTCAAACTATGACAA 1284
 DB 444 ACATTGCTCACATGAAGCCCAAGGAGCAGCTGTAGAGTGGACTTCAACACAAATGTGGA 385
 QY 1285 GCGAAGATTTACTGAGGGCTTTGAGAACAGTCATTACCGATTCCCTCTTATAAAGAGAATG 1344
 DB 384 GTACAGACTTGTGCTGAATGCAATTGAAGAGAGTAATTATGATCCTTCATATAAGAGAATG 325
 QY 1345 CTATCAGATTATCAAGAATTCACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGTCT 1404
 DB 324 TTATGAAATTTATCAAGAATTCACCATGATCAACCAAGTGAAG-CCCTGGATCGAGCAGTCT 266
 QY 1405 TCTGATCGAGTTTGTCTATGCGCCCAAGAGGAGCCACAGCCTGGGATCAGCTGCCCATG 1464
 DB 265 TCTGATTTGAATTTGTCTATGCGCCCAAGAGGAGCTTAAACACCTTCGGGTTGCGAGCCAG 206
 QY 1465 ACCTCAGCTGGTTCCAGCAGTACTCTATAGATGATGGTGGTCTGCTGCTGACCTGTGTGG 1524
 DB 205 ACCTCAGCTGGTTCCAGTACTCTTTGGATGATGATGGTGGTCTGCTGCTGCTGTGTGG 146
 QY 1525 CAACCTGCTATATCTTGTTCACAAATGTTTTTATTTTCTGTCGCAAAATTTAAATAA 1584
 DB 145 CACTGTGATATTTATGTCACAAATGTTGCTGTTTGTCTGTTTCTGGAAGTTTGTCTAGAA 86
 QY 1585 CTAGAAAGATAGAAAGAGGGAATAGATGATCTTTT 1617
 DB 85 AAGCAAGAGGGAAGAAAATGATTAGTATATC 53

RESULT 11
 BG249053
 LOCUS
 DEFINITION 602361828F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4470199 5',
 mRNA sequence.
 ACCESSION BG249053
 VERSION BG249053.1 GI:12758869
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10287 row: h column: 08
 High quality sequence stop: 661.
 Location/Qualifiers

FEATURES

```
source
1. .671
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4470199"
/clone_lib="NIH_MGC_89"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      206 a 123 c 153 g 189 t
ORIGIN
Query Match      28.6%; Score 467.4; DB 12; Length 671;
Best Local Similarity 94.2%; Pred. No. 5.6e-103;
Matches 606; Conservative 0; Mismatches 21; Indels 16; Gaps 11;
Qy 193 TGACTCACTCAAGCCTTCGTTAATGACTACAGAGCCCTTCGCAATTGAAATTTGAGG 252
Db 11 TGACTCACTCAAGCCTTCGTTAA-TGACTACAGAGCCCTTCGCAATGAA--TTGAGG 67
Qy 253 TGGTCCATATGCCACAGACAGACAGAGAAATGAATATTTGTTGACCTAGCTCTGA 312
Db 68 TGGTCCATATGCCACAGACAGACAGAGAAATGAATA-TTGTTCACCTAGCTCTGA 126
Qy 313 ATGTCTTCCAGGCTTATCAACCTGGCAATCAGTTATAAATTAATGATTTTTTGTG 372
Db 127 ATGTCTTCCAGGCTTATCAACCTGGCAATCAGTTATAAATTAATGATTTTTTGTG 186
Qy 373 AAATAAGAGGAACCTTAAAAATGATGTGTGAGAGCTTATCTACAATCAGACGCTTATGA 432
Db 187 AAATAAGAGGAACCTTAAAAATGATGTGTGAGAGCTTATCTACAATCAGACGCTTATGA 246
Qy 433 AGAAGCTACAGAGAAACCACTACGATGTAATGCTTTATAGACCCCTGTAATCCCTGTGGAG 492
Db 247 AGAAGCTACAGAGAAACCACTACGATGTAATGCTTTATAGACCCCTGTAATCCCTGTGGAG 306
Qy 493 ACCTGATGGCGAGTGTGCTGCAGTCCCTTTGTGCTCACACTTAGAATTTCT--GTAGGA 551
Db 307 ACCTGATGGCGAGTGTGCTGCAGTCCCTTTGTGCTCACACTTAGAATTTCTGGTAGGA 366
Qy 552 GGCAATATGGAGCGAAGCT--GTGGAACTTCCAGCTCCA--CTTTCCTATGACCTGTGC 609
Db 367 GGCAATATGGAGCGAAGCTGTGGAACTTCCAGCTCCAAGCTTCTCTATGACCTGTGC 426
Qy 610 CTATGACAGGAGTAACACAGAGAA-TGACCTTTCTGGAAAGAGT---AAAAAATTCATG 665
Db 427 CTATGACAGGAGTAACACAGAGAGTGACCTTTCTGGAAAGAGTCAACACATTCATG 486
Qy 666 CTTTCAGTTTGTTCCTACTTCTGGATTACGAGTATCAGTATCATTTTGGGAAGAGTTT 725
Db 487 CTTTCAGTTTGTTCCTACTTCTGGATTACGAGTATCAGTATCATTTTGGGGAGAGTG 546
Qy 726 T--ATAGTAAGGCATTAGGAAGGCCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGA 783
Db 547 TGAGTAGTAAGGCAATTAGGAAGGCCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGA 606
Qy 784 TATGG-CTAATACCAACATATTGGGATTTGCAATTTCTCTCAAC 825
Db 607 TATGGCCTAATACCAACATAT--GGGATTTGAATTTCTCTCAAC 647

RESULT 12
BI331895 LOCUS      977 bp mRNA linear EST 30-JUL-2001
DEFINITION 602982580F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5135491 5', mRNA sequence.
ACCESSION BI331895
VERSION BI331895.1 GI:15016552
KEYWORDS EST.
SOURCE house mouse.
```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 977)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM11332 row: p column: 20
High quality sequence stop: 873.
Location/Qualifiers
1. .977
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5135491"
/clone_lib="NCI_CGAP_L19"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 267 a 218 c 222 g 270 t
ORIGIN

Query Match 28.4%; Score 464.4; DB 13; Length 977;
Best Local Similarity 72.5%; Pred. No. 3.5e-102;
Matches 657; Conservative 0; Mismatches 241; Indels 8; Gaps 4;
Qy 633 ATGACCTTTCTGGAAGAGTAAAAAATTCATGCTTTCAGTTCCTTTCCTCCTCTCGGATT 692
Db 1 ATGACATTTGCGAGAAAGGTCAGAAATATGTTGCAGGTGTTGCTTTTGGACTTTTGGTTT 60
Qy 693 CAGGATTACGACTATCATTTTGGGAAGAGTTTATAGTAGGATTTAGGAGGCCCCACT 752
Db 61 CAACATTTAAACAGAAATCTCGGAATCAGTTTACAGTGTGTTCTAGGAGAGCTTACA 120
Qy 753 ACATTATGTGAGACTGTGGGAAAGCTGAGATATGGCTTAATACGAACATATTTGGGATT 812
Db 121 ACATTACTGAGATGATGGGAGGAGCAGACATATGGCTCGTTCGAACCTTCTGGGACTG 180
Qy 813 GAATTTCTCAACCATACCAACCTTAACCTTTGAGTTTGTGGAGGATTCAGTCTGTAACCT 872
Db 181 AAATTTCTCACCTTTCTTGCCTAAATTTGACTTTGTTGGAGGACTCCATTTGTAACCA 240
Qy 873 GCCAAAGCTTTGCCCTAAGGAAATGGAATTTTTCAGAGTTTCAGGGAGAGATGGTATT 932
Db 241 GCCAAACCACTGCCCTAAGGAAATGGAATTTTTCAGAGCTCTGGAGACATCGGTGTG 300
Qy 933 GTGGTGTTCCTCTGGGTCTACTGTTTCAAAATTTGTACAGAAAGAAAGGCTTAATATCAT 992
Db 301 GTGGTGTTCCTCTGGGATCAATGTTTAAACATTTAAAGAAAGAAAGGCCAATGTAGTT 360
Qy 993 GCTTCAGCCCTTGCCCGAGATCCCAAGAGGTGTTATGGAGGTACAAAGGAAAAAACCA 1052
Db 361 GCTTCTGCTCTTGCCCGAGATTCACAGAGGTTCTGTGGAGATTTGATGGTAAGAAACCA 420
Qy 1053 TCCACATTAGGAGCCCACTACTCGGCTGTATGATTTGGATACCCAGATGATCTTCTGGT 1112
Db 421 GACACCTTAGGATCAACACTCGGCTCTACAAGTGGATCCCGGAGATGATCTTCTGGT 480
Qy 1113 CATCCCAAAACCAAGCTTTTATCACATCATGCTGGAATGAATGGAATCATGTAAGACTATT 1172
Db 481 CATCAAAACCAAGCTTTTATAGCTCATGCTGGAACCAATGCAATCATGAGCGATC 540

1173	QY	TACCATGGGGTCCCTATGGTGGGAGTTCCCATATTTGGTGATCAGCTTGATACATAGCT	1232
541	Db	TACCATGGCATCTCTATTGTTGGTATTCCCTGTGTTGGGATCAACCTGATAATAATAC	600
1233	QY	CACATGAAGGCCAAGGAGGACGCTGTAGAATAAATCAAACTTGAACAACCGAGAT	1292
601	Db	CACATAGTAGCAAGGAGCAGCAGCTGTTAGAGTGGACTTTCATACATGCTCAACTAGACAC	660
1293	QY	TTACTGAGGCGCTTGAGAACACAGTCATTTACCGATTCTCTTTATAAGAGAATGCTATGAGA	1352
661	Db	CTTCTCACTGCCCTGGAAGACTGTCATTAATGACCTTCCTTATAAGAGAATGCCATGAGA	720
1353	QY	TTATCAAGAATTCACCATGATCA-ACCTGTTAAAGCCCCCTAGATCGAGCAGTCTTCTGGAT	1411
721	Db	TTATCCAGAATCCACCATGACCAGACCATTGAAGCCCTTTGGACCGAGCTGTCTCTGGGA	780
1412	QY	CGAGTTTGT---CATGGGCCACAAGGAGGCCAA---GCACCTGCGATAGCTAGCTGCCCATG-	1464
781	Db	TCGACGTATGTCCATCGCAACAAGGGAACCAAGCAACCTTCGCCCAAGCTCTCATGGA	840
1465	QY	ACCTACCTGGTTCCAGCACTACTCTATAGATGTGATGGGTTCTGCTGACCTGTGTGG	1524
841	Db	ACCTTACCTGGTTCCAGTACCACCTCTCTGGATGGTGACTGGGTCTATAGGCCCTGGGTGG	900
1525	QY	CAACTG 1530	
901	Db	TACCTG 906	

[illegible]

DEFINITION	AGNCOURT_8804675 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6337257
ACCESSION	5' mRNA sequence.
VERSION	BQ942862
KEYWORDS	BQ942862.1 GI:22358160
SOURCE	EST.
ORGANISM	house mouse. Mus musculus. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 938)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-i@mail.nih.gov Tissue Procurement: Susan L. Sullivan, PhD. cDNA Library Preparation: Resgen, Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 10:35:27 ; Search time 63.8246 Seconds
(without alignments)
7860.974 Million cell updates/sec

Title: US-09-981-353-165
Perfect score: 1636
Sequence: 1 gatcagtggtgagggaact.....ccaaattcaagaagacacgtg 1636

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	775.6	47.4	2107	4	US-09-180-852-1
2	574	35.1	1413	4	US-09-813-918-1
3	329	20.1	2339	5	PCT-US92-00282-2
4	319.8	19.5	2336	5	PCT-US92-00282-1
5	145.6	8.9	391	4	US-09-370-838-21
6	97.4	6.0	1561	5	PCT-US92-00282-25
7	83.8	5.1	783	5	PCT-US92-00282-22
8	72.2	4.4	1008	5	PCT-US92-00282-10
9	62.4	3.8	1448	5	PCT-US92-00282-16
10	60	3.7	1066	5	PCT-US92-00282-14
11	55.4	3.4	1190	5	PCT-US92-00282-18
12	52.6	3.2	1197	5	PCT-US92-00282-20
13	50.6	3.1	1800	6	5180581-1
14	50.6	3.1	2793	1	US-08-281-916-5
15	50.6	3.1	2793	2	US-08-460-725-7
16	48.8	3.0	1669	3	US-08-522-421-1
17	44	2.7	1627	3	US-09-106-464-1
18	41.4	2.5	51	5	PCT-US92-00282-37
19	38.8	2.4	1607	2	US-08-797-226-1
20	38.6	2.4	2590	4	US-09-623-062-2
21	38.6	2.4	5194	1	US-09-599-652-1
22	38.6	2.4	5194	2	US-08-642-846-1
23	38.6	2.4	5194	4	US-09-264-604-1
24	38.6	2.4	5333	4	US-09-623-062-1
25	38.2	2.3	2025	4	US-08-942-0128-23
26	37.2	2.3	1219	5	PCT-US92-00282-8
27					

C 28	36.2	2.2	2450	3	US-08-617-860B-9	Sequence 9, Appl 1
C 29	36	2.2	4342	4	US-09-338-907-107	Sequence 107, App
C 30	36	2.2	4342	4	US-09-218-207-107	Sequence 107, App
C 31	36	2.2	4582	4	US-09-338-907-118	Sequence 118, App
C 32	36	2.2	4582	4	US-09-218-207-118	Sequence 118, App
C 33	36	2.2	4686	4	US-09-338-907-117	Sequence 117, App
C 34	36	2.2	4686	4	US-09-218-207-117	Sequence 117, App
C 35	36	2.2	4875	4	US-09-338-907-114	Sequence 114, App
C 36	36	2.2	4875	4	US-09-218-207-114	Sequence 114, App
C 37	36	2.2	4958	4	US-09-338-907-116	Sequence 116, App
C 38	36	2.2	4958	4	US-09-218-207-116	Sequence 116, App
C 39	36	2.2	4986	4	US-09-338-907-121	Sequence 121, App
C 40	36	2.2	4986	4	US-09-218-207-121	Sequence 121, App
C 41	36	2.2	5020	4	US-09-338-907-120	Sequence 120, App
C 42	36	2.2	5020	4	US-09-218-207-120	Sequence 120, App
C 43	36	2.2	5044	4	US-09-338-907-115	Sequence 115, App
C 44	36	2.2	5044	4	US-09-218-207-115	Sequence 115, App
C 45	36	2.2	5057	4	US-09-338-907-123	Sequence 123, App

ALIGNMENTS

RESULT 1
US-09-180-852-1
; Sequence 1, Application US/09180852
; Patent No. 6287834
; GENERAL INFORMATION:
; APPLICANT: BELANGER, Alain
; APPLICANT: HUM, Dean W.
; APPLICANT: BEAULIEU, Martin
; APPLICANT: LEVESQUE, Eric
; TITLE OF INVENTION: CHARACTERIZATION AND USE OF AN ISOLATED URIDINE
; TITLE OF INVENTION: DIHOSPHO-GLUCURONOSYLTRANSFERASE
; FILE REFERENCE: 1259-449
; CURRENT APPLICATION NUMBER: US/09/180,852
; CURRENT FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: PCT/CA97/00328
; EARLIER FILING DATE: 1997-05-16
; EARLIER APPLICATION NUMBER: US 08/649,319
; EARLIER FILING DATE: 1996-05-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(1644)
US-09-180-852-1

Query Match	47.4%	Score 775.6;	DB 4;	Length 2107;
Best Local Similarity	68.6%	Pred. No. 4.5e-216;		
Matches 1104;	Conservative	0;	Mismatches 494;	Indels 12; Gaps 2;
QY	24	ATCATGAGTCTGCACAGTCAGCTTGGTATTTCTGCTCTCGACGCTTCTCTGT	---	GTT 80
Db	46	ACCAGGATGCTCTGAATGATGTCAGTCTTCTGCTGATGCGAGCTCAGTTGTTACTTT		
QY	81	GGCTGTGGATCTTGTGGAAAGTCCTGTGTGGCCCTGTGACATGAGCCATGGCTTAAT		
Db	106	AGCTCTGGAGTCTTGGAAAGTCTGTGTGGCCACAGATACACCCATGGATTAAT		
QY	141	GTCAAGGTCACTTAGAAGAGCTCATAGTCAGAGGCCATGAGGTAAACAGTATTGACTCAC		
Db	166	ATGAAGACATCTTGGAGAGCTTGTTCAGAGGGGTCATGAGGTGATTGTGTGACATCT		
QY	201	TCAAAGCTTCTGTTAATTGACTACAGGAAGCCCTTCTGATTTGAAATTTGAGGTGGTCCAT		
Db	226	TCGGCTTCTATTCTTGTCAATGCCAGTAATCATCTGCTATTAAATTAAGATTTATCT		
QY	261	ATGCCACAGGACAGACAGAAATGAAATATTT-----GTTGACCTAGCTCTG		

Db 286 ACATCTTTAACTAAAAATGATTTGGAGATTTTTTATGAAAAATGTCGATAGTGACA 345
 QY 312 AATGCTCTGCGAGGCTTATCAACCTGGCAATCAGTTATATAAATAAATGATTTTGTGTT 371
 Db 346 TATAGTATTTCAAAAAATACATTTTGGTCATATTTTTCACAACTACAAAGAAATGTTGG 405
 QY 372 GAAATAGAGGAACCTTTTAAAAATGATGTGTGAGAGCTTTATCTACAACTACAGAGCTTATG 431
 Db 406 GAATATTTCTGACATATATAAAGCTGTGTGAAGATGCGAGTTTGAACAAGAACTTATG 465
 QY 432 AGAAGCTACAGAAACCACTACAGATGTAATGCTTATAGACCTGTGATTCCTCTGGGA 491
 Db 466 AGAAAACTACAAGATCAAAATTTGATGTCCTTCTGCGAGATCCGTTAATCCCTGTGGT 525
 QY 492 GACCTGATGCTCAGTTGCTGAGTCCCTTTTGTGCTCACACTTAGAATTTCTGTAGGA 551
 Db 526 GAGCTGCTGCTGAACACTACTTAACATACCTTTCTGTACAGTCTCCGCTTCTCTGTGGC 585
 QY 552 GGCATATATGAGCGAAGCTGTGGGAACTTCCAGCTCCACTTTCCCTATGTACCTGTGCT 611
 Db 586 TACACAGTTGAGAAGAAATGCTGGAGGATTTCTGTCTCCCTTCTTCTATGTACCTGTGTT 645
 QY 612 ATCAGAGGACTACAGACAGATGACCTTTCTGGAAGAGTAATAAATTCATGCTTTCA 671
 Db 646 ATGTCAGAATTAAGTATCAAAATGATTTTTCATGGAGAGGATAAAAAATATGATATATG 705
 QY 672 GTTTTCTGCTCAGTTCTGAGTTCAGGATAGGACTATCATTTTGGGAAGATTTTATAGT 731
 Db 706 CTTATTTTGGCTTTGTTTGAAGCATATGATCTGAAGAGTGGGACCAAGTTTATAGT 765
 QY 732 AAGGCATATAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGATATGGCTA 791
 Db 766 GAAGTTCTAGGAAGACCCACTACATTTATGTGACAAATGGGAAAGCTGAAATGTGCTC 825
 QY 792 ATACGAACATATTTGGATTTTGAATTTCTTCACACCATACCAACCTAACCTTGTGTTGTT 851
 Db 826 ATTCGAACCTATTTGGATTTTGAATTTCTCGGCCATTTCTTACCAATGTGATTTGTT 885
 QY 852 GGAGGATGTCACGTAAACCTGCGAAAGCTTTCCCTTAAGGAAATGGAATTTTGTCCAG 911
 Db 886 GGAGGACTCAGTGTAAACCGCAACCCCTGCTTAAAGAAATGGAAGAGTTGTGCGAG 945
 QY 912 AGTTCAGGGAAGATGATTTGTGTTTCTCTGGGTCACCTGTTTCAAAATGTTACA 971
 Db 946 AGCTCTGGGAAAAATGATTTGGTGTGTTTCTCTGGGTCGATGATCAGTAACATGTC 1005
 QY 972 GAAGAAAGCTTAATCATTTGCTTCCAGCCCTTGGCCAGATCCCAAGAGGTGTTATGG 1031
 Db 1006 GAAGAAAGTGCACCAATGATTTGATCAGCCCTTGGCCAGATCCCAAGAGGTGTTATGG 1065
 QY 1032 AGGTACAAAGGAAAAAACCACATCCATTTAGGAGCAATCTCGGCTGTATGATTTGATA 1091
 Db 1066 AGATTTGATGGCAAGCAAAATCTTTAGTGTCCAACTACGCTGTATAGTGGTTA 1125
 QY 1092 CCCAGAAATGATCTTCTGCTATCCCAAAACAAAGCTTTTATCACTCATGCTGGAATG 1151
 Db 1126 CCCAGAAATGATCTTCTGCTATCCCAAAACAAAGCTTTTATCACTCATGCTGGAATG 1185
 QY 1152 AATGGGATCTAGCAACTATTTACCATGGGTCCTTATGGTGGAGTTCCTCATTTGTT 1211
 Db 1186 AATGGCATCTATGAGCGCATCTACCATGGGATCCCTTATGGTGGGATTCCTTGTGGG 1245
 QY 1212 GATCAGCTTGATAACATAGCTACATGAAGGCAAGAGGAGCAGCTGTAGAAATTAACCTTC 1271
 Db 1246 GATCAACATGATAACATTTGCTACATGAAGCAAGGAGGAGCAGCCCTCAGTGTGGACATC 1305
 QY 1272 AAAACTATCAGACGGAAGATTTTACTGAGGCTTTTGAGAACAGTATACCCGATTCCTCT 1331
 Db 1306 AGGACCATGTCAGTAGAGATTTGCTCAATGCATTTCAAGTCAAGTCAATTAATGACCCATC 1365
 QY 1332 TATAAGAGANTGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTGTAAAGCCCTTA 1391

Db 1366 TATAAGAGAAATATCATGAATAATTCAGAAATTCATCATGATCAACGGGTGAAGCCCTG 1425
 QY 1392 GATCGAGCAGCTCTTCTGGATCGAGTTTGTCTATCGGCACAAAGAGCAACGACCTGCGA 1451
 Db 1426 GATCGAGCAGCTCTTCTGGATCGAGTTTGTCTATCGGCATAAAGAGCAACGACCTCGG 1485
 QY 1452 TCAGCTGCCCATGACCTCAGCTGCTGCTCCAGCAGTCTCTATAGATGTGATGGTTCCTG 1511
 Db 1486 GTCCGAGCCCAACCACTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1545
 QY 1512 CTGACCTGTGTGCAACTGCTATATCTCTTCCACAAATGTTTATTTTCTGTCAT 1571
 Db 1546 CTGCGCTGCTGCACTATGATATTTATCATCAAAATGTTGCTGTTTGTGTTCCGA 1605
 QY 1572 AAATTTAAATAAATAAGAGATAGAAAGAGGAAATAGATAGATCTTTCCAA 1621
 Db 1606 AAGCTTCCCAACAGGAAGAAAGAGGATAGTTATATATCAAA 1655

RESULT 2
 US-09-813-918-1
 : Sequence 1, Application US/09813918
 : Patent No. 6383789
 : GENERAL INFORMATION:
 : APPLICANT: WEBSTER, Marlon et al.
 : TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
 : TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
 : TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
 : TITLE OF INVENTION: AND USES THEREOF
 : FILE REFERENCE: CL001175
 : CURRENT APPLICATION NUMBER: US/09/813,918
 : CURRENT FILING DATE: 2001-03-22
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 1413
 : TYPE: DNA
 : ORGANISM: Human
 US-09-813-918-1

Query Match 35.1%; Score 574; DB 4; Length 1413;
 Best Local Similarity 78.4%; Pred. No. 2e-157;
 Matches 688; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 743 AAGGCCCACTACATTTATGTGAGACTGTGGGAAAGCTGAGATATGCTAATACGAACATA 802
 Db 520 ACGACCCACTACCTTATTTGAGCAATGGGAAAGCTGACATATGCTTATGCAAAACC 579
 QY 803 TTGGGATTTGAATTTCTCCTCAACCATACCACTTACCTTGTGTTGTTGGAGATGCA 862
 Db 580 CTGGAGTTTTCATTTCTCTCATCTCTTACCAACGTTGATTTTGTGGAGATTTCCA 639
 QY 863 CTCTAAACCTGCCAAAGCTTTGCTTAAGGAAATGGAATTTTGTCCAGAGTTCCAGGGA 922
 Db 640 CTCAAAACCTGCCAAACCCCTACTTAAGGAAATGGAGAGTTTGTACAGAGCTCTGGAGA 699
 QY 923 AGATGTTATTTGTTGTTTCTCTGGGTCACCTGTTTCAAAATGTTTACAGAAAGAGGC 982
 Db 700 AAATGTTGTTGTTGTTTCTCTGGGTCAGTGATGAAGTAACATGACACAGAAAGGGC 759
 QY 983 TAATATCATTTGCTTACGCTTCCGAGATCCGAGATGCTTATGAGGTTACAAAGG 1042
 Db 760 CAATGTAATTTCAACAGCCCTTCCAGGATTCACAAAGGTTCTGTGAGATTTGACGG 819
 QY 1043 AAAAAAACCATCCACATTAGGAGCCCAATCTCGCTGTATGATTTGATACCCAGATGA 1102
 Db 820 GAATAAACCAAGATGCTTCTAGGCTCTCAATCTCGGCTGTACAAGTGGATACCCAGATGA 879
 QY 1103 TCTTCTTGTGTCATCCCAAAACCAAGCTTTTATCACTCATGTTGGAATGGAATGGATCTA 1162
 Db 880 CCTTCTAGGTCATCCCAAAACCAAGAGCTTTTAACTCATGTTGGAGCCCAATGGCATCTA 939
 QY 1163 TGAAGCTATTTACATGGGGTCCCTATGTTGGGAGTTCCCATATTTGGTGATCAGCTTGA 1222

```

940 TGAGGCAATCTACCATGGGATCCCATGGTGGGCAATCCATGTTTTTGGATCAACCTGA 999
1223 TAACATAGCTACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAACTATGAC 1282
1000 TAACATTTCTCACATGAAGGCCAAAGGAGCAGCTGTAGATTGGACITCAACAAATGTC 1059
1283 AAGCGAAGATTACTGAGGGCTTTGAGAACAGTCAATACCGATTCTCTTTATAAGAGAA 1342
1060 GAGTACAGACCTGCTGAATGACACTGAAGACAGTAATTAATGATCCTTTATATAAGAGAA 1119
1343 TGCTATGAGATTATCAAGAATTACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGT 1402
1120 TATTATGAATTTATCAAGAATTCAACATGATCAACCAAGTAAAGCCCTGGATCGAGCAGT 1179
1403 CTTCTGATCAGTTTGTGATGCGCCCAAGAGGAGCAAGCAGCTGGGATCAGTGCCCA 1462
1180 CTTCTGATTTGAATTTGTGATGCGCCCAAGAGGAGCAACACCTTCGAGTTGCAGGCCA 1239
1463 TGACCTCACCTGGTTCAGGCACTACTCTATAGATGTGATGGGTCTCTGCTGACCTGTGT 1522
1240 TGACCTCACCTGGTTCAGGATCAGTCTCTTTGGATGTGATGGGTCTCTGCTGACCTGTGT 1299
1523 GGCAACTGCTATATCTGTGTTCAAAATGTTTTTATTTCTGTTCAAAATTTAATAA 1582
1300 GGCAACTGCTATATTTATCATCATCAAAAGTTTGTGTTGTTTCTGGAAGTTTGTCTAG 1359
1583 AACTAGAAAGTAAAGAGGAGGATAGATCTTTTCCAA 1620
1360 AAAAGGAGAGGAGGAGGAGGAGGATAGTATGTCGGA 1397

```

RESULT 3

```

PCT-US92-00282-2
: Sequence 2, Application PC/TUS9200282
: GENERAL INFORMATION:
: APPLICANT: OWENS, IDA S.
: APPLICANT: RITTER, JOSEPH K.
: TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
: TITLE OF INVENTION: THEREIN.
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN DABY & CUSHMAN
: STREET: 1615 L STREET, N.W.
: CITY: WASHINGTON
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20036-5601
: COMPUTER TYPE: Floppy disk
: MEDIUM TYPE: PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/00282
: FILING DATE: 19920110
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: SCOTT, WATSON T.
: REGISTRATION NUMBER: 26581
: REFERENCE/DOCKET NUMBER: 91532-PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-861-3000
: TELEFAX: 202-822-0944
: TELETYPE: 6714627 CUSH
: INFORMATION FOR SEQ. ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2339 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
PCT-US92-00282-2

```

```

Query Match      20.1%; Score 329; DB 5; Length 2339;
Best Local Similarity 52.3%; Pred. No. 5.9e-86;
Matches 797; Conservative 0; Mismatches 710; Indels 12; Gaps 3;
QY 57 CTGCTCTCTGAGCTCTTCTGTGTGGCTGTGGATCTGTGGGAAAGTCTCTGTGGTGGCC 116
Db 46 CTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 111
QY 117 TGTGACATGAGCCATTGGCTTAATGTCAGAGTCTTCTAGAGAGAGTCTATAGTGAGAGC 176
Db 106 ACTGATGGAGCCCTCTGGCTCAGCATGCGGAGGCTTGCAGGAGCTCTCATGCCAGAGC 165
QY 177 CATGAGGTAACTGATTTGACTCCTCAAGCCCTTCTGTAATTTGACTACAGGAGCTTCT 236
Db 166 CACCAGCGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 225
QY 237 GCATTGAATTTGAGGTGGTCCATATGCCACAGGAGCAGAAACAGAGAAATGAATATTT 296
Db 226 ACCCTGACAGCCCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 279
QY 297 GTTGACCTAGCTCTGAATGCTCTTGCAGGCTTATCAACCTGGCAATCAGTTATATAAAT 356
Db 280 CTGGGCTACATCAAGGGTCTTCTGAAACAGAACATCTTCTGAAGAGATATTTAGAAAT 339
QY 357 AATGATTTTTTGTGTAATTAAGAGAACTTTAAATATGATCTGTGAGAGCTTTATCTAC 416
Db 340 ATGGCAATTTATCAACAATGATCTTTGGCCCTTCTATAGTCTGTGTGGAGCTACTGCAT 399
QY 417 AATCAGAGCTTATGAAGAAGCTACAGGAAACCACTACGATTAATGCTTTATAGACCT 476
Db 400 AATGAGCCCTGATCAGCACCCTGAATGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 459
QY 477 GTGATTCCTCTGGAGACCTGATGGCTGTGCTGTGCTGAGTCCCTTTTGTGCTCACACT 536
Db 460 GTTAACCTCTCGGGGGGCTGCTGGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
QY 537 AGAATTTCTGTAGGAGCAATATCGAGGAGAGCTGTGGAACTTCAGCTCCACTTTCC 596
Db 517 TGGAGGTACATTCCTCTGAGTCTTAAAGGACACAGTGTCCAAATCTCTCTCTCTCT 576
QY 597 TATGTACCTGTGCTCTATGACAGGACTAACAGAGAGTAACTTCTCTGGAAGAGTAA 656
Db 577 TATATCTCTAGTTACTAACGACCAATTCAGACCATGACATCTCTGCAAGAGGTCAG 636
QY 657 AATTCATGCTTTCTGAGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 716
Db 637 AAC---ATGCTCTACCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 693
QY 717 GAAGAGTTTTATAGTAAGGCAATGAGGAGCCCTACATTTATGAGAGCTGTGGGAAA 776
Db 694 GCAAGTCTTGCCTCTGAGCTTTTTCAGAGAGAGTGTGAGTGTGATGCTGTGCTGCTAT 753
QY 777 GCTGAGATATGCTATACGAACATATGAGGATTTTGAATTTTCTCAACCATACCAACT 836
Db 754 GCATCGGTGTGCTGCTGAGGAGCTTTGATGAGTCTACCTACCCAGGCGGATCATGCC 813
QY 837 AACTTTGAGTTTGTGGAGGATTTGAGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTG 896
Db 814 AACATGCTCTTCTGAGGAGTCACTTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAG 873
QY 897 GAAAATTTTGTGAGGAGTTTTCAGGAGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTG 956
Db 874 GAAGCCTACATTAATGCTTCTGGAGAACATGGAATTTGTTTCTCTCTGCTGCTGCTG 933
QY 957 TTTCAAAATGTTACAGAGAAAGGCTAATATCATTTGCTTTCAGCCTTTGCCAGATCCCA 1016
Db 934 GTCTCAGAAATTCAGAGAGAGGAGGAGTATGGAATTTGCTGCTGCTGCTGCTGCTG 993
QY 1017 CAGAGAGGTTTATGGAGGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1076
Db 994 CAGAGAGCTCTGTGGCGGTACTGGAACCGGACCATCGAATCTTTGGGAGAACAGGATA 1053

```

Qy	1077	CTGTATGATTTGGATACCCAGAATGATCTCTTTGGTCATCCCAAAACCAAGCTTTTATC	1136
Db	1054	TTTTTTAAAGTGGGTACCCAAAACGATCTGCTTGGTCACCCGATGACCGTGCTTTATC	1113
Qy	1137	ACTCATGGTGAATGAATGGGATCTATGAAGCTATTATACCATGGGTCCCATATGGTGGGA	1196
Db	1114	ACCCATGCTGGTTCCCATGGTGTATTGAAGACATATGCATCGCGTTCCCATGGTGATG	1173
Qy	1197	GTTTCCCATATTTTGGTGATCAGCTTGATTAACATAGCTTCACATGAAGGCCAAAGGACGAGCT	1256
Db	1174	ATGCCCTTGTTTGGTGATCAGATGGGACAAATGCAAAAGCGCATGAGAGACTTAAGGAGCTGGA	1233
Qy	1257	GTAGAAATAAATTCACAAACTATGACAAGCGAAGATTTACTGAGGCGTTTGACAAACAGTC	1316
Db	1234	GTGACCTGAACTGTTCTGGAAATGACTTCTGAAGATTTTGAATAATGCTCTAAAGGAGTC	1293
Qy	1317	ATTACCGATTCCTCTTATAAAGAGAATGCTATGAGATTTATCAAGAATTCAACCATGATCAA	1376
Db	1294	ATCAATGACAAAAAGTTACAAGGAGAAACATCAGGCGCTCTCCAGCCTCTACAAAGGACCGC	1353
Qy	1377	CGTGTAAAGCCCTAGATCGAGCAGCTTCTGCGATCGAGTTTGTCTATGGCCCAACAAGGA	1436
Db	1354	CCGGTGGAGCCGCTGGACCTTGGCCGTGTTCTGGGTGGAGTTTGTGATGAGGCCACAAGGC	1413
Qy	1437	GCCAAGCACCTGCGATCAGCTGCCCATGACCTCACCTCGTGTCCAGCAGCTACTCTATAGAT	1496
Db	1414	GCGCCACACCTGGCCGCCGAGCCAGCAGCCTCACCTGGTACCAGTACCATTCTCTGGAC	1473
Qy	1497	GTGATTTGGTTCCTGCTGACCTGTGTGGCAACCTGCTATTTCTGTGTTCAAAAATCTTTT	1556
Db	1474	GTGATTTGGTTTCCTCTTTGGCCGTGCTGACACATGGCCCTTCATACCTTTAAATGTGT	1533
Qy	1557	TTATTTTTCCTGTCAAAAAT	1575
Db	1534	GCTTATGGCTACCGGAAAT	1552

RESULT 4

```

PCT-US92-00282-1
; Sequence 1, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; APPLICANT: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2336 base pairs

```

QY 1055 CACATTAGGAGCAATACCTCGCTGTATGATGGATACCCAGAAATGATCTCTTGGTCA 1114
Db 1029 GAATCTGGCAACACAGATACCTTGAAGTGGCTACCCAAAACGATCTGCTGGTCA 1088
QY 1115 TCCAAAACCAAGCTTTTATCACATCATGGTGAATGAATGGATCTATGAAGCTATTTA 1174
Db 1089 CCGGATGACCCGCTGCTTTATCACCATGCTGGTCCCATGGTGTATGAAGCATATG 1148
QY 1175 CCATGGGTCCCTATGGTGGGATGCTCCCATATTTGGTGATCAGCTTGATAACATAGCTCA 1234
Db 1149 CAATGGGTTCCTATGGTGGGATGATGCTCCCTTGTGGTGATCAGATGGACAATGCAAGCG 1208
QY 1235 CATGAAGCCAAAGGAGCAGCTGTAGAAATTAACCTTCAAACTATGCAAGCGAAGATTT 1294
Db 1209 CATGAGACTAAGGAGCTGGAGTGACCTGAATGTTCTGGAATGACTTCTGAAGATTT 1268
QY 1295 ACTGAGGCTTTGAGAACAGTCATACCGATTCTCTTATAAGAGAATGCTATGAGATT 1354
Db 1269 AGAAATGCTCTAAAGCAGTCATCAATGACAAAGTTACAAGGAGAATCATCAGGCGCT 1328
QY 1355 ATCAAGAAATCACCATGATCAACCTGTAAAGCCCTAGATCGAGCAGTCTCTGGATCGA 1414
Db 1329 CTCCAGCCTTCAAGAGACCCCGGTGGAGCCCTGGACCTGGCCGTGTTCTGGGTGA 1388
QY 1415 GTTGTCTATGCGCCACAAAGGACCAAGCAGCTGCGATCAGCTGCCCATGACCTCAGCTG 1474
Db 1389 GTTTGTGATGAGGACAAAGGCGCGCCACACCTTGGCGCCGAGCCAGCAGCTCAGCTG 1448
QY 1475 GTTCAGCAGTCTATAGATGATGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534
Db 1449 GTACAGTACCATCTTCTGACGAGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1508
QY 1535 ATTCTGTTCCAAAATGTTTATTTTCTCTGTCAAAAT 1575
Db 1509 CTTTCATCCTTTAAATGTTGCTGCTTATGCTACCGGAAT 1549

RESULT 5
US-09-370-838-21
; Sequence 21, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT FILING DATE: 1999-08-09
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-21

Query Match 8.9%; Score 145.6; DB 4; Length 391;
Best Local Similarity 61.2%; Pred. No. 6.4e-33;
Matches 235; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 925 ATGCTATGCTGTTTCTCTGGGGTCACTGTTTCAAAATGTTACAGAGAAAAGGCTA 984
Db 1 ATGGAATGCTGTTTCTCTTTGGGATCAATGGTCTCAGAAATCCAGAGAGAAAGCTG 60
QY 985 ATATCATGCTTCAGCCCTTCCAGATCCCAAGAGAGGTTTATGAGGTTACAAAGAA 1044
Db 61 TGGCATGCTGATGCTTTGGGCAAAATCCCTCAGACAGTCTCTGTCGGGTACACTGGA 120

QY 1045 AAAACCATCCATATAGGAGCAATACCTCGCTGTATGATGGATACCCAGAAATGATC 1104
Db 121 CCGAGCATCGAATCTTGGCAACACACGATACCTTGTAGTGGCTACCCAAAACGATC 180
QY 1105 TTCTTGTCTATCCCAAAACCAAGCTTTTATCACTCATCTGTTGAATGAATGGATCTATG 1164
Db 181 TGCTTGTCAACCAATGACCCGTCCTTATACCCATGCTAGTTCCTCCATGGTGTAAATG 240
QY 1165 AAGCTATTTACATGGGGTCCCTATGTTGGGAGTTCCCATATTTGGTGATCAGCTTATA 1224
Db 241 AAAGCATATGCAATGCGCTTCCCATGGTATGATACCTTATTTGGTGATCAGATGACA 300
QY 1225 ACATAGCTCACATGAAGGCAAGGAGCAGCTGTAGAAATTAACCTTCAAACTATGACAA 1284
Db 301 ATGCAAGCGCAGGAGACTAAGGAGCTGGAGTGGAGTGGCTGATGTTCTGGAGATGACTT 360
QY 1285 GCGAAGATTACTGAGGCTTTGA 1308
Db 361 CTGAGATCTAGAAGATGCTCTGA 384

RESULT 6

PCT-US92-00282-25
; Sequence 25, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; APPLICANT: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UCT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1561 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US92-00282-25

Query Match 6.0%; Score 97.4; DB 5; Length 1561;
Best Local Similarity 70.1%; Pred. No. 1.4e-18;
Matches 131; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1331 TTATAAGAGAAATGCTATGAGATTATCAAGAAATTCACCATGATCAACCTGTAAGCCCT 1390
Db 271 TTACAAGGAGAACATCATGCGCTCTCCAGCTTCAAGAGACCCCGCGGTGGAGCCGCT 330
QY 1391 AGATCAGCAGTCTTCTGATCGAGTTTGTCTATGCGCCACAAAGAGGACCAAGCAGCTCG 1450

Db 331 GGACCTGGCCGTGTTCTGGGTGGAGTTGTGATGAGCACAAGGCGCGCACACCTGCG 390
Qy 1451 ATGAGCTGCCCATGACCTACCTGGTTCAGCAGCACTACTCTATAGATGTGATGGTTCCT 1510
Db 391 CCCGCGAGCCACAGCAGCTACCTGGTGTACCACTACCATTCCTTGGAGCGTGTGTTTCCT 450
Qy 1511 GCTGACC 1517
Db 451 CTTGGCC 457

RESULT 7

PCT-US92-00282-22
; Sequence 22, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US92-00282-22

Query Match 5.1%; Score 83.8; DB 5; Length 783;
Best Local Similarity 61.9%; Pred. No. 9.1e-15;
Matches 133; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
Qy 1110 GGTATCCCAACCAAGCTTTTATCACTCATGCTGGTGAATGATGGATCTATGAAGCT 1169
Db 471 GGTACCCGATGACCCCTGCTTTATCACCATGCTGTTCCCATGCTTTATGNAAGC 530
Qy 1170 ATTTACCATGGGTCCTATGTTGGGAGTTCCTCATATTTGGTGATCAGCTTGATAACATA 1229
Db 531 ATATGCAATGGCTTCCCATGGTGATGATGCCCTTGTGTTGGTGTATGATGGACAATGCA 590
Qy 1230 GCTCAGATGAAGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAACTATGACAAGCGAA 1289
Db 591 AAGGCATGAGACTAAGGAGCTGGAGTGACCCCTGAATGTTCTCGGAAATGACCTCTGAA 650
Qy 1290 GATTACTAGGGCTTTGAGACAGTCAATACCA 1324
Db 651 GATTAGAAAATGCTCTAAAAGCAGTCAATGA 685

RESULT 8

PCT-US92-00282-10
; Sequence 10, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; APPLICANT: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1008 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US92-00282-10

Query Match 4.4%; Score 72.2; DB 5; Length 1008;
Best Local Similarity 45.5%; Pred. No. 2.5e-11;
Matches 380; Conservative 0; Mismatches 443; Indels 12; Gaps 3;
Qy 57 CTGCTCCTGCAGCTCTTCTGTGTTGGCTGTGGATTCTGTGGGAAGCTCTGTGTGGTGGCCC 116
Db 157 CTGCTGCTTCTCTCTCAGTGTCCAGCCCTGGGCTGAGAGTGGGAAGGTGCTGTGGTGGCCC 216
Qy 117 TGTGACATGAGCCATTCGCTTAATGTCAAGTCTATTCTAGAAGAGCTCATAGTCAGAGGC 176
Db 217 ACTGATGGCAGCCACTGGCTCAGCATGGGAGGCCCTTGGGGAGCCCTCCATGCCAGAGGC 276
Qy 177 CATGAGTAACAGTATTGACTCACTCAAAGCTTCCTGTTAATTTGACTACAGGAAGCCCTTCT 236
Db 277 CACGAGTGGTGGTCTCACCTCGAGGTGAATATGTATACATCAAGAAGAGAACTTTTTC 336
Qy 237 GCATTGAAATTTGAGGTGGTCCATATGCCACAGGACAGACAGAAAGAAATGAAATATTT 296
Db 337 ACCCTGACAACGTATGCCATTTTCATGGACCCAGGAC-----GAATTTGATCGCCTTTTG 390
Qy 297 GTTGACCTAGCTCTGAATGTCTTCCAGGCTTATCAACCTGGCAATCAGTTATAAATA 356
Db 391 CTGGGTCACTCAATCGTTCTTTTGAACACAGAACATCTTCTGATGAAATTTTCTAGAAGA 450
Qy 357 AATGATTTTTTTGTTGAAATAAGAGGAACATTTTAAATAATGATGTGTGAGAGCTTTTATCTAC 416
Db 451 ATGCAATTTATGAACAATATGCTTTTGTATCATAGTCTGTGTGGAGCTACTGCAT 510
Qy 417 AATCAGAGCTTTATGAAGAAGCTTACAGGAACCAACTACGATGTAATGCTTTATAGACCT 476

: SEQUENCE CHARACTERISTICS:

LENGTH: 1066 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

```

;
; TOPOLOGY: linear
;

```

MOLECULE TYPE: CDNA

PCT-US92-00282-14

Query Match 3.78: Score 60: DB 5: Length 1066:

Query Match 3.7%; Score 00; Length 1000;
Best Local Similarity 44.9%; Pred. No. 9.1e-08;
Matches 360; Conservative 0; Mismatches 430; Indels 12; Gaps 3;

QV 57 CTGCTCCTGCAGCTCTTCTCTGTGTTGGCTGTGGATTCTGTGGAAAGTCCTGGTGTGGCCC 116

Db 176 CTGCTGCTTCTCCTCAGTGTCAGCCCTGGCTGAGAGTGGAAAGGTGTTGGTGGTGCCC 235

QY 117 TGTGACATGAGCCATTGGCTTAATGTCAAGGTCATTCTAGAAGAGCTCATAGTGAGAGGC 176

Db 236 ATTCATGGCAGCCACTGGCTCAGCATGCGGAGGTCTTGGGGAGCTCCATGCCAGAGGC 295

177 CATGAGGTACAGTATTGACTCACTCAAAGCCTTCGTTAATTGACTACAGGAAGCCTTCT 236

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

D6 296 CACCAGGCAGTGGTCCCTCACCCTCAGAGGTGAATATGCACATCAAGAAGAGAACTTTTC 355

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

237 GCATTTGAAATTTTGAGGTGGTCCATATGCCACAGGACAGAACAGAAATATTT 296

100

DD 336 ACCCTGACAACTATGCCATTTTCGTGGACCCAGGATGAAATTTGATCGCCATG-----TG 409

OV 297 GTTGACCTAGCTCTGAATGTCCTTGGCAGGCTTATCAACCTGGCAATCAGCTTATAAAATTAA 356

QJ QIQAACCTAGCCTCAATGCTTCCAGGCCTATCATACCCTGGCATTCAGTTATAAATAIA 350

ZS7 QIQAACCTAGCCTCAATGCTTCCAGGCCTATCATACCCTGGCATTCAGTTATAAATAIA 350

Db 410 CTGGGCCACACTCAACTGTACTTTGAAACAGAACATTTTCTGAAGAAATTTTTCAGAAAGT 469

1

QY 357 AATGATTTTTTGTGTTGAAATAAGAGGAACTTTAAAAATGATGTGTGAGAGCTTTATCTAC 416

Db 470 ATGGCAATGTTGAACAATATGCTCTTTGGCTATCATAGGCTTGTGTGGAGCTACTACAT 529

Qy 417 AATCAGACGCTTATGAAGAAGCTACAGGAAACCAACTACGATGTAATGCTTATAGACCCCT 476

A vertical ruler with markings from 0 to 10 cm. The markings are in millimeters, with major ticks every centimeter and minor ticks every millimeter. The ruler is oriented vertically with the 0 mark at the top and the 10 cm mark at the bottom.

D_b 530 AATGAGGCCCTGATCAGGCACCTGAATGCTACTTCCTTTGATGTGGTTTAAACAGACCCC 589

Qy 477 GTGATTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCCTTTTGCTCACA

RESULT 11

PCT-US92-00282-18

; Sequence 18, Application PC/TUS9200282

; GENERAL INFORMATION:

```

QY 575 GAAACTTCCAGCTCCACTTTCTATGACCTGTGCTATGACAGGACTAACAGACAAAT 634
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 CCAGTGGCCCAACCACTTCTCTAGTGGCCAGCCCTCTCTCTCTCATTCAGATACAT 695
QY 635 CACCTTTCTGGAAGAGTAAATAATCAATGCTTTTCAAGTTTGTTCACACTT 685
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 GACCTTCTCGAGCGGTGAGAACATGCTCATTCGCTTTTTCACAGAACTT 746

```

RESULT 12

```

PCT-US92-00282-20
; Sequence 20, Application PC/TUS9200282
; GENERAL INFORMATION:
; APPLICANT: OWENS, IDA S.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282.
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 20:
; LENGTH: 513 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US92-00282-20

```

```

Query Match 3.4%; Score 55; DB 5; Length 513;
Best Local Similarity 61.5%; Pred. No. 1.8e-06;
Matches 88; Conservativeness 0; Mismatches 55; Indels 0; Gaps 0;
QY 882 TTGCTTAAGGAATGGAATAATTTCTCCAGAGTTTCAGGGGAGAGTGTATTGGTGT 941
Db 222 TGGCTCTAGGAATTTGAAGCCCTACATTAATGCTTCTGGAGAACATGGAAATTTG 281
QY 942 TCTCTGGGTGCTACTGTTTCAAAATGTTACAGAAGAAAGGCTAATATCATTCGCT 1001
Db 282 TCTTTGGGATCAATGGTCTCAGAAATTCAGAGAAGAAAGCTATGCGCAATTCGCT 341
QY 1002 CTTGCCAGATCCACAGAAGT 1024
Db 342 TTGGGCAAAATCCCTCAGACAGT 364

```

RESULT 13

```

PCT-US92-00282-12
; Sequence 12, Application PC/TUS9200282
; GENERAL INFORMATION:

```

```

; APPLICANT: OWENS, IDA S.
; APPLICANT: RITTER, JOSEPH K.
; TITLE OF INVENTION: THE GENETIC LOCUS UGT1 AND A MUTATION
; TITLE OF INVENTION: THEREIN.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/00282
; FILING DATE: 19920110
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26581
; REFERENCE/DOCKET NUMBER: 91532-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 12:
; LENGTH: 1197 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US92-00282-12

```

```

Query Match 3.2%; Score 52.6; DB 5; Length 1197;
Best Local Similarity 47.1%; Pred. No. 1.4e-05;
Matches 232; Conservativeness 0; Mismatches 254; Indels 7; Gaps 2;
QY 399 TGTGAGAGCTTTATCTACAATACAGCGCTTATGAAGAGCTACAGGAAACCACTAGAT 458
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 612 TGTGTGGAGCTACTGCATAATGAGCCCTGATCAGGCACCTGATGCTACTTCTTTGAT 671
QY 459 GTAATGCTTATAGACCTGTGATTCCCTGTGGAGACCTGATGGCTGCTGTCAGTC 518
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 672 GTGTTTTTAACAGACACCCCGTTAACCTCTGTGGGCGGTGCTGCTAAGTACCTGTG 731
QY 519 CCTTTTGTGCTCACACTTAGAATTTCTGTAGGAGCAATATGAGCGAAGCTGTGGAAA 578
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 732 CTGCTCTGTTTTTTTGGAGGTACATTCATGTGAC----TTAGACTTTTAAGGCCATACA 787
QY 579 CTTCACCTCCACTTTCTCTATGACCTGTGCTTATGACAGGACTTAACAGACAGAAATGACC 638
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 GTGTCAAATCTCTCTCTATATTTCTTAAGTACTTAACGACCAATTCAGACCACATGACA 847
QY 639 TTTCTGAAAGAGTAAATAATTCATGCTTTTCAGTTTGTTCACCTCTGCTGATTCAGGAT 698
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 TTCTGCAAGAGGTCAAGAAC---ATGCTCTACCCCTCTGGCCCTGTCTTACATTTGCCAT 904
QY 699 TAGCACTATCATTTTTTGGGAAGAGTTTATAGTAAGGCAATTAGGAAGGCCCACTACATTA 758
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ACTTTTCTGCCCTTATGCAAGCTTTCCTCTGAGCTTTTCAGAGAGAGGTGTCAGTG 964
QY 759 TGTGAGACTGTGGGAAAGAGCTGAGATATGCTTAATACGAACATATTGGGATTTGAATTT 818
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 GTGGATCTTGTGTCAGTATGCTATCCGTGTGCTTCCGAGGGGACTTTGTGATGGACTAC 1024
QY 819 CCTCAACCACTACCAAGCTTAACCTTTGAGTTTGTGGAGGATTCAGCTGTAACCTGCCAAA 878
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 CCCAGGCCGATCATGCCCAACATGGTCTTCAATTTGGGGGCGATCACTGTGCCCAACGGGAG 1084

```

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 29-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
TELEX: 49617824
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
US-08-281-916-5

Query Match 3.1%; Score 50.6; DB 1; Length 2793;
Best Local Similarity 49.1%; Pred. No. 8.1e-05;
Matches 134; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 1036 ACAAGGAAAAAACCATCCACATTAGGAGCCAACTACTCGGCTGTATGATTGGATACCCC 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1143 ACGAAGTAGTAAAAACATAACGTTGCCGCCCAACGTAAATACGCAAAATTTGTTTAATC 1202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1096 AGAATGATCTCTTGTCATCCCAACCAAGCTTTTATCACTCATGGTGAATGAATG 1155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1203 AAGCGCGCTGTCGCTATAAAAAAATGGCGGCTTTATACGCAAGCGGACTACAAAT 1262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1156 GGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGAGTTCCCATATTTGGTGATC 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1263 CGAGCGACGAGGCGCTTGGAAAGCGGGATACCCATGGTGTCTGCCCCATGATGGCGGACC 1322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1216 AGCTTGATACATAGCTCATCATGAAGCCCAAGGACGACGTGTAGAAATAAACTTCAAAA 1275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1323 AGTTTATACCATCGGCACAATTTACAGCAACTCGGCGTAGCCCGCGCTTGGACACTGTTA 1382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1276 CTATGACAAGCGGAAGATTTTACTGAGGGCTTTGA 1308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1383 CCGTTCCAGCGATCACTACTAGTGGCGGATAA 1415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: July 16, 2003, 14:58:55
Job time : 67.8246 secs

```

```

QY      879 GCTTGGCCTAAGC 891
        |   |   |||
Db     1085 .CCACTATCTCAGG 1097

RESULT 14
5180581-1
; Patent No. 5180581
; APPLICANT: MILLER, LOIS K.; O'REILLY, DAVID R.
; TITLE OF INVENTION: BIOLOGICAL INSECT CONTROL AGENTS AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,952
; FILING DATE: 29-JUN-1989
; SEQ ID NO:1:
; LENGTH: 1800
5180581-1

Query Match          3.1%; Score 50.6; DB 6; Length 1800;
Best Local Similarity 49.1%; Pred. No. 6.5e-05;
Matches 134; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY      1036 ACAAGAGAAAAAACCATCCACATAGGAGCAAACTCGGTGTATGATTGGATACC CC 1095
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1143 ACGAAGTAGTAAAAAACATAAGCTGCCCGCCCAACGTAATCACGCNAATTTGGTTAATC 1202

QY      1096 AGAATGATCTTCCTGGTCAATCCAAAACCAAAGCTTTTATCACTCAATGCTGGAATGAATG 1155
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1203 AACGGCCCGTGCTGGCTCATAAAAAAATGGCGCGTTTTATTACGCAAGCGGACTACAAT 1262

QY      1156 GGATCTATCAAGCTATTTACCATGGGTCCCTATGTTGGGAGTTCCTCATATTTGGTGATC 1215
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1263 CGAGCGACGAGCGCTTGGAAAGCCGGATACCCATGGTGCTGCCCCATGATGGCGGACC 1322

QY      1216 AGCTTGATAACATAGCTCACTGAAGCCAAAGGACGACCTGTAGAAATAAACTTCAAAA 1275
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1323 AGTTTTACCATCGGCACAAATATACAGCAACTCGGGCTAGCCGCGCTTGGACACTGTGTA 1382

QY      1276 CTATGACAGCGCAAGATTTACTGAGGCGCTTGA 1308
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1383 CCGTTTTCCAGCGATCAACTACTAGTGGCGGATAA 1415

RESULT 15
US-08-281-916-5
; Sequence 5, Application US/08281916
; Patent No. 5682897
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Black, Bruce C.
; APPLICANT: Dierks, Peter M.
; APPLICANT: Fleming, Nancy C.
; APPLICANT: Ahmed, Fakhruddin
; TITLE OF INVENTION: Insect Viruses, Sequences, Insecticidal
; METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; STREET: Greenlee and Winner, P.C.
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/281,916
; FILING DATE: 27-JUL-1994

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 13:28:07 ; Search time 215,273 seconds

(without alignments)

12133.023 Million cell updates/sec

Title: US-09-981-353-165

Perfect score: 1636

Sequence: 1 gatcagtggtgagggaact.....ccaaattcaagaagacctg 1636

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1114047 segs, 798260406 residues

Total number of hits satisfying chosen parameters: 2228094

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBOMB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1636	100.0	1636	9	US-09-981-353-165
2	1636	100.0	2866	9	US-09-981-353-33
3	1636	100.0	2974	9	US-10-174-590-521
4	1636	100.0	2974	9	US-10-176-758-521
5	1636	100.0	2974	9	US-10-175-737-521
6	1636	100.0	2974	9	US-10-173-706-521
7	1636	100.0	2974	9	US-10-175-738-521
8	1636	100.0	2974	9	US-10-175-752-521
9	1636	100.0	2974	9	US-10-176-482-521
10	1636	100.0	2974	9	US-10-176-757-521
11	1636	100.0	2974	9	US-10-176-913-521
12	1636	100.0	2974	9	US-10-180-552-521
13	1636	100.0	2974	9	US-10-180-557-521
14	1636	100.0	2974	9	US-10-173-700-521
15	1636	100.0	2974	9	US-10-174-572-521
16	1636	100.0	2974	9	US-10-174-579-521
17	1636	100.0	2974	9	US-10-174-582-521
18	1636	100.0	2974	9	US-10-174-588-521
19	1636	100.0	2974	9	US-10-175-739-521

20	1636	100.0	2974	9	US-10-175-740-521	Sequence 521, App
21	1636	100.0	2974	9	US-10-175-743-521	Sequence 521, App
22	1636	100.0	2974	9	US-10-176-488-521	Sequence 521, App
23	1636	100.0	2974	9	US-10-176-492-521	Sequence 521, App
24	1636	100.0	2974	9	US-10-176-747-521	Sequence 521, App
25	1636	100.0	2974	9	US-10-176-750-521	Sequence 521, App
26	1636	100.0	2974	9	US-10-176-985-521	Sequence 521, App
27	1636	100.0	2974	9	US-10-176-987-521	Sequence 521, App
28	1636	100.0	2974	9	US-10-176-991-521	Sequence 521, App
29	1636	100.0	2974	9	US-10-176-992-521	Sequence 521, App
30	1636	100.0	2974	9	US-10-176-993-521	Sequence 521, App
31	1636	100.0	2974	9	US-10-184-658-521	Sequence 521, App
32	1636	100.0	2974	9	US-10-173-695-521	Sequence 521, App
33	1636	100.0	2974	9	US-10-173-697-521	Sequence 521, App
34	1636	100.0	2974	9	US-10-173-705-521	Sequence 521, App
35	1636	100.0	2974	9	US-10-174-576-521	Sequence 521, App
36	1636	100.0	2974	9	US-10-174-585-521	Sequence 521, App
37	1636	100.0	2974	9	US-10-174-586-521	Sequence 521, App
38	1636	100.0	2974	9	US-10-175-747-521	Sequence 521, App
39	1636	100.0	2974	9	US-10-176-481-521	Sequence 521, App
40	1636	100.0	2974	9	US-10-176-485-521	Sequence 521, App
41	1636	100.0	2974	9	US-10-176-487-521	Sequence 521, App
42	1636	100.0	2974	9	US-10-176-493-521	Sequence 521, App
43	1636	100.0	2974	9	US-10-176-756-521	Sequence 521, App
44	1636	100.0	2974	9	US-10-176-911-521	Sequence 521, App
45	1636	100.0	2974	9	US-10-176-919-521	Sequence 521, App

ALIGNMENTS

RESULT 1

US-09-981-353-165
; Sequence 165, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 165
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2434655CBI
US-09-981-353-165

Query Match	100.0%	Score 1636;	DB 9;	Length 1636;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1636;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
Qy	1	GATCAGTCTGTGAGGAACTGCCATCATGAGTCTGACAAAGTCAGCTTTGGTATTTCG 60		
Db	1	GATCAGTCTGTGAGGAACTGCCATCATGAGTCTGACAAAGTCAGCTTTGGTATTTCG 60		
Qy	61	TCCTGCAGCTCTCTGCTGTGGCTGTGGATTCTGTGGGAAAGTCTCTGTGGCCCTCTG 120		
Db	61	TCCTGCAGCTCTCTGCTGTGGCTGTGGATTCTGTGGGAAAGTCTCTGTGGCCCTCTG 120		
Qy	121	ACATGAGCCATGGCTTAAATGTCAAGGTCTTCTAGAGAGCTCATAGTGAGAGCCATG 180		
Db	121	ACATGAGCCATGGCTTAAATGTCAAGGTCTTCTAGAGAGCTCATAGTGAGAGCCATG 180		
Qy	181	AGGTAACAGTATTGACTCACTCAAGCCCTTCGTTAAATTTGACTACAGGAGCCCTTCTGCAT 240		
Db	181	AGGTAACAGTATTGACTCACTCAAGCCCTTCGTTAAATTTGACTACAGGAGCCCTTCTGCAT 240		

Qy 241 TGAATTTAGGTGGTCCATATGCCACAGGACAGACAGAGAAAATGAAATTTGTTG 300
Db
Qy 241 TGAATTTAGGTGGTCCATATGCCACAGGACAGACAGAGAAAATGAAATTTGTTG 300
Db
Qy 301 ACCTAGCTCTGAATGCTTGGCCAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATG 360
Db
Qy 301 ACCTAGCTCTGAATGCTTGGCCAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATG 360
Db
Qy 361 ATTTTGTGTTGAATTAAGAGAACTTTAAAATGATGCTGGAGAGCTTTATCTACAAATC 420
Db
Qy 361 ATTTTGTGTTGAATTAAGAGAACTTTAAAATGATGCTGGAGAGCTTTATCTACAAATC 420
Db
Qy 421 AGACGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTTATAGACCTGTGA 480
Db
Qy 421 AGACGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTTATAGACCTGTGA 480
Db
Qy 481 TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGGAGTCCCTTTTGTGCTCACACTTAGAA 540
Db
Qy 481 TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGGAGTCCCTTTTGTGCTCACACTTAGAA 540
Db
Qy 541 TTTCTGTAGGAGCAATATGAGGCAAGCTGTGGGAACTTCCAGCTCCACTTTTCTTATG 600
Db
Qy 541 TTTCTGTAGGAGCAATATGAGGCAAGCTGTGGGAACTTCCAGCTCCACTTTTCTTATG 600
Db
Qy 601 TACCTGTGCTTATGACAGGACTTAACAGACAGAAATGACCTTTCTGGAAGAGTAAAAAT 660
Db
Qy 601 TACCTGTGCTTATGACAGGACTTAACAGACAGAAATGACCTTTCTGGAAGAGTAAAAAT 660
Db
Qy 661 CAATGCTTTCAGTTTGTTCACCTTCTGATTCAGGATACGACTATCAATTTTGGGAAG 720
Db
Qy 661 CAATGCTTTCAGTTTGTTCACCTTCTGATTCAGGATACGACTATCAATTTTGGGAAG 720
Db
Qy 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTTATGAGACTGTGGGAAAGCTG 780
Db
Qy 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTTATGAGACTGTGGGAAAGCTG 780
Db
Qy 781 AGATATGGCTAATACGAACATATTTGGGATTTTGAATTTCTTCAACCACTAACTAACT 840
Db
Qy 781 AGATATGGCTAATACGAACATATTTGGGATTTTGAATTTCTTCAACCACTAACTAACT 840
Db
Qy 841 TTGAGTTTGTGGAGATTCACATGTAACCTGCCAAGCTTTTCCCTAAGGAATGAAA 900
Db
Qy 841 TTGAGTTTGTGGAGATTCACATGTAACCTGCCAAGCTTTTCCCTAAGGAATGAAA 900
Db
Qy 901 ATTTTGTCCAGAGTTACAGGGAAGATGATTTGTTGTTTCTCTGGGCTCACTGTTTC 960
Db
Qy 901 ATTTTGTCCAGAGTTACAGGGAAGATGATTTGTTGTTTCTCTGGGCTCACTGTTTC 960
Db
Qy 961 AAAATGTTACAGAAAGGCTAATATCATTTGCTTCCAGCTTCCAGATCCACAGA 1020
Db
Qy 961 AAAATGTTACAGAAAGGCTAATATCATTTGCTTCCAGCTTCCAGATCCACAGA 1020
Db
Qy 1021 AGGTGTTATGAGGTACAAAGGAAAAAACCACATAGGAGCAATACCTCGGCTGT 1080
Db
Qy 1021 AGGTGTTATGAGGTACAAAGGAAAAAACCACATAGGAGCAATACCTCGGCTGT 1080
Db
Qy 1081 ATGATTTGGATACCCAGAAATGATCTTCTGGTATCCCAACCAAGCTTTTATCACTC 1140
Db
Qy 1081 ATGATTTGGATACCCAGAAATGATCTTCTGGTATCCCAACCAAGCTTTTATCACTC 1140
Db
Qy 1141 ATGTTGGAATGGAATGGGATCTATCAAGCTATTTACCATGGGTCCTTATGGTGGAGTTC 1200
Db
Qy 1141 ATGTTGGAATGGAATGGGATCTATCAAGCTATTTACCATGGGTCCTTATGGTGGAGTTC 1200
Db
Qy 1201 CCATATTTGGTGATCAGCTTGATAACATAGCTACATGAAGGCCAAAGGAGCAGCTGTAG 1260
Db
Qy 1261 AAATAAACTTCAAAACTATGACAAAGCGAAGATTTTACTGAGGGCTTTTGAACACAGTCA 1320
Db
Qy 1261 AAATAAACTTCAAAACTATGACAAAGCGAAGATTTTACTGAGGGCTTTTGAACACAGTCA 1320
Db
Qy 1321 CCGATTCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380

Db 1321 CCGATTCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Qy 1381 TAAAGCCCTAGATCGAGCAGCTTCTTGGATCGAGTTTGTCTATCGCCACAAAGAGCCA 1440
Db 1381 TAAAGCCCTAGATCGAGCAGCTTCTTGGATCGAGTTTGTCTATCGCCACAAAGAGCCA 1440
Qy 1441 AGCACCTGCGATCAGCTGACCTGACCTGCTGTTCCAGCAGCTACTCTATAGATGTGA 1500
Db 1441 AGCACCTGCGATCAGCTGACCTGACCTGACCTGCTGTTCCAGCAGCTACTCTATAGATGTGA 1500
Qy 1501 TTGGTTCTCTGCTGACCTGTGTGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
Db 1501 TTGGTTCTCTGCTGACCTGTGTGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
Qy 1561 TTTCTCTGCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGAATAGATCTTTCCAA 1620
Db 1561 TTTCTCTGCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGAATAGATCTTTCCAA 1620
Qy 1621 ATTCAAGAAAGACCTG 1636
Db 1621 ATTCAAGAAAGACCTG 1636

RESULT 2
US-09-981-353-33
; Sequence 33, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 2966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 997080.1
US-09-981-353-33

Query Match 100.0%; Score 1636; DB 9; Length 2966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGCAAGCTCAGCTTTGGTATTCTGCG 60
Db 4 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGCAAGCTCAGCTTTGGTATTCTGCG 63
Qy 61 TCCTGCAGCTCTTCTGTTGGCTGTGGATTTCTGGGAAAGTCTTGGTGTGGCCCTGTG 120
Db 64 TCCTGCAGCTCTTCTGTTGGCTGTGGATTTCTGGGAAAGTCTTGGTGTGGCCCTGTG 123
Qy 121 ACATGAGCCATTTGGCTTAAATGCTCAAGGTCTTCTAGAGAGCTCATAGTGAAGGCCATG 180
Db 124 ACATGAGCCATTTGGCTTAAATGCTCAAGGTCTTCTAGAGAGCTCATAGTGAAGGCCATG 183
Qy 181 AGGTAACAGATTTGACTCAGCTCAAAAGCTTCTGTTAAATGACTACAGGAAGCTTCTGCAT 240
Db 184 AGGTAACAGATTTGACTCAGCTCAAAAGCTTCTGTTAAATGACTACAGGAAGCTTCTGCAT 243
Qy 241 TGAATTTGAGGTGGTCCATATGCCAGGACAGACAGAGAAAATGAAATATTTGTTG 300
Db 244 TGAATTTGAGGTGGTCCATATGCCAGGACAGACAGAGAAAATGAAATATTTGTTG 303
Qy 301 ACCTAGCTCTGAATGCTTGGCAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATG 360
Db 304 ACCTAGCTCTGAATGCTTGGCAGGCTTATCAACCTGGCAATCAGTTATATAAATTAATG 363

QY 361 ATTTTGTGTTGAAATGAAGAGGACTTTAAATGATGTGTGAGAGCTTTATCTACAATC 420
Db 364 ATTTTGTGTTGAAATGAAGAGGACTTTAAATGATGTGTGAGAGCTTTATCTACAATC 423
QY 421 AGACGCTTATGAAGAGCTACAGGAACCACTACGATGTAATGCTTATAGACCCGTGTA 480
Db 424 AGACGCTTATGAAGAGCTACAGGAACCACTACGATGTAATGCTTATAGACCCGTGTA 483
QY 481 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTGAGTCCCTTTGTGCTCACACTTAGAA 540
Db 484 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTGAGTCCCTTTGTGCTCACACTTAGAA 543
QY 541 TTCTGTAGGAGCAATATGAGCGAAGCTGTGGGAACTTCAGACTCCACTTTCTCTATG 600
Db 544 TTCTGTAGGAGCAATATGAGCGAAGCTGTGGGAACTTCAGACTCCACTTTCTCTATG 603
QY 601 TACTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAT 660
Db 604 TACTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAT 663
QY 661 CAATGCTTTCAGTTTGTCCACTCTGGAATCAGGATACGACTATCATTTTGGGAAG 720
Db 664 CAATGCTTTCAGTTTGTCCACTCTGGAATCAGGATACGACTATCATTTTGGGAAG 723
QY 721 AGTTTATAGTAAGGCAATAGGAAGCCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
Db 724 AGTTTATAGTAAGGCAATAGGAAGCCCACTACATTTATGTGAGACTGTGGGAAAGCTG 783
QY 781 AGATATGGCTAATACGAACATATGGGATTTTGAATTTCTCAACCATACCAACCTAACT 840
Db 784 AGATATGGCTAATACGAACATATGGGATTTTGAATTTCTCAACCATACCAACCTAACT 843
QY 841 TTCAGTTTGTGAGAGTTCACCTGTAAACCTGCCAAGCTTTCCTTAGGAAATGGA 900
Db 844 TTCAGTTTGTGAGAGTTCACCTGTAAACCTGCCAAGCTTTCCTTAGGAAATGGA 903
QY 901 ATTTTCCAGAGTTTCAGGGAAGATGATTTGTTGTTGTTTCTCTGCGGTCACTTTTC 960
Db 904 ATTTTCCAGAGTTTCAGGGAAGATGATTTGTTGTTGTTTCTCTGCGGTCACTTTTC 963
QY 961 AATATGTACAGAAAGGCTAATATCATTTGCTTCAGCCCTTGCCAGATCCCAACA 1020
Db 964 AATATGTACAGAAAGGCTAATATCATTTGCTTCAGCCCTTGCCAGATCCCAACA 1023
QY 1021 AGGTGTTATGAGGTACAAAGGAAAAAACCATCCACATTAGAGCCATACCTCGCTGT 1080
Db 1024 AGGTGTTATGAGGTACAAAGGAAAAAACCATCCACATTAGAGCCATACCTCGCTGT 1083
QY 1081 ATGATTTGATACCCAGAAATGATCTTTTGGTATCCCAAAACCAAGCTTTTATCACATC 1140
Db 1084 ATGATTTGATACCCAGAAATGATCTTTTGGTATCCCAAAACCAAGCTTTTATCACATC 1143
QY 1141 ATGGTGAATGAATGGGATCTATGAAGCTATTTACCATGGGTCCCTATGTTGGGATTC 1200
Db 1144 ATGGTGAATGAATGGGATCTATGAAGCTATTTACCATGGGTCCCTATGTTGGGATTC 1203
QY 1201 CCATATTTGGTATGATGATGATATACATAGCTACATGAAGCCCAAGGAGCAGCTGTAG 1260
Db 1204 CCATATTTGGTATGATGATGATATACATAGCTACATGAAGCCCAAGGAGCAGCTGTAG 1263
QY 1261 AATAAATTTCAAACTATGACAAGCGAAGATTTTACTGAGGGCTTTGAGAAGCAGTCATTA 1320
Db 1264 AATAAATTTCAAACTATGACAAGCGAAGATTTTACTGAGGGCTTTGAGAAGCAGTCATTA 1323
QY 1321 CCGATTTCTTTATGAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTG 1380
Db 1324 CCGATTTCTTTATGAAGAGATGCTATGAGATTTATCAAGATTCACCATGATCAACCTG 1383
QY 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCTATGCGCCACAAGAGCCA 1440
Db 1384 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCTATGCGCCACAAGAGCCA 1443

QY 1441 AGACACCTGGATCAGCTGCCCATGACCTCACCTGGTTCAGCACTACTCTATAGATGTGA 1500
Db 1444 AGACACCTGGATCAGCTGCCCATGACCTCACCTGGTTCAGCACTACTCTATAGATGTGA 1503
QY 1501 TTGGGTTCTGTGCTGACCTGTGGCACTGCTATATTTCTGTTCAAAAATGTTTTTAT 1560
Db 1504 TTGGGTTCTGTGCTGACCTGTGGCACTGCTATATTTCTGTTCAAAAATGTTTTTAT 1563
QY 1561 TTCTCTGTCAAAAATTTAATAAACTAGAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1620
Db 1564 TTCTCTGTCAAAAATTTAATAAACTAGAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1623
QY 1621 ATCAAGAAGACCTG 1636
Db 1624 ATCAAGAAGACCTG 1639

RESULT 3
US-10-174-590-521
; Sequence 521, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-590-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAAGTCAGCTTTGGTATTTCGTC 60
Db 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAAGTCAGCTTTGGTATTTCGTC 60
QY 61 TCCTCAGCTCTTCGTGTTGGTCTGTTGGTCTGTTGGTCTGTTGGTCTGTTGGTCTGTTGG 120
Db 61 TCCTCAGCTCTTCGTGTTGGTCTGTTGGTCTGTTGGTCTGTTGGTCTGTTGGTCTGTTGG 120
QY 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCATTCTAGAAGAGCTCATAGTGAGAGCCATG 180
Db 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCATTCTAGAAGAGCTCATAGTGAGAGCCATG 180
QY 181 AGGTAAACAGTATTGACTCAGTCAAGAGCTTTGTTAAATGACTACAGGAGCCCTTCGTGAT 240
Db 181 AGGTAAACAGTATTGACTCAGTCAAGAGCTTTGTTAAATGACTACAGGAGCCCTTCGTGAT 240
QY 241 TGAATTTTGGGTGGTGGTCCATATGCCACAGGACAGAAAGAAATGAAATATTTGTTG 300
Db 241 TGAATTTTGGGTGGTGGTCCATATGCCACAGGACAGAAAGAAATGAAATATTTGTTG 300
QY 301 ACCTAGCTCTGAATGCTTTGCCAGGCTTATCAACCTGCAATCAGTTATATAAATTAATG 360
Db 301 ACCTAGCTCTGAATGCTTTGCCAGGCTTATCAACCTGCAATCAGTTATATAAATTAATG 360

Qy	361	ATTTT	TTGTTGAAAT	TAAGAGGAAC	TTTAAAA	TGATGCTG	GAGAGCT	TTTATCT	ACAATC	420	
Db	361	ATTTT	TTGTTGAAAT	TAAGAGGAAC	TTTAAAA	TGATGCTG	GAGAGCT	TTTATCT	ACAATC	420	
Qy	421	AGACG	CTTATGAAGA	AGCTACAG	GAACCAAC	TACGAT	GTAAATG	CTTATAG	ACCCCTGTGA	480	
Db	421	AGACG	CTTATGAAGA	AGCTACAG	GAACCAAC	TACGAT	GTAAATG	CTTATAG	ACCCCTGTGA	480	
Qy	481	TTCCCT	GTGAGAC	CTGATGG	CTGAGTTG	CTGCAG	TCCTTTT	TGCTCAC	ACTTAGAA	540	
Db	481	TTCCCT	GTGAGAC	CTGATGG	CTGAGTTG	CTGCAG	TCCTTTT	TGCTCAC	ACTTAGAA	540	
Qy	541	TTTCTG	TAGAGCA	ATATG	GAGGAAG	CTGTGG	GAACCTT	CCAGCT	CCACTTTCCTATG	600	
Db	541	TTTCTG	TAGAGCA	ATATG	GAGGAAG	CTGTGG	GAACCTT	CCAGCT	CCACTTTCCTATG	600	
Qy	601	TACCTG	TGCTATG	CAGGACT	AACGAC	AGAATG	ACCTTTT	CTGGAA	GAGTAAAAAAT	660	
Db	601	TACCTG	TGCTATG	CAGGACT	AACGAC	AGAATG	ACCTTTT	CTGGAA	GAGTAAAAAAT	660	
Qy	661	CAATG	CTTTCAG	TTTGTTC	CACTTCTG	GAATCAG	GAATAC	GAATAT	CAATTTTGGGAAG	720	
Db	661	CAATG	CTTTCAG	TTTGTTC	CACTTCTG	GAATCAG	GAATAC	GAATAT	CAATTTTGGGAAG	720	
Qy	721	AGTTT	TATAGT	AAGCAAT	AGGAAG	CCCACTAC	ATTATG	TGAGACT	GTGGGAAAAGCTG	780	
Db	721	AGTTT	TATAGT	AAGCAAT	AGGAAG	CCCACTAC	ATTATG	TGAGACT	GTGGGAAAAGCTG	780	
Qy	781	AGATG	GGCTPAAT	CAGAAC	ATATG	GGATTTT	GAAATTC	TCTCA	ACCATACCAACCTAACT	840	
Db	781	AGATG	GGCTPAAT	CAGAAC	ATATG	GGATTTT	GAAATTC	TCTCA	ACCATACCAACCTAACT	840	
Qy	841	TTGAG	TTGTTG	GAGGATTC	CACTGTAA	ACCTGCC	AAAGCTTTG	CCCTA	AGGAATGGAAA	900	
Db	841	TTGAG	TTGTTG	GAGGATTC	CACTGTAA	ACCTGCC	AAAGCTTTG	CCCTA	AGGAATGGAAA	900	
Qy	901	ATTTG	TCCAGACT	TCAGGG	AAAGATG	GGTATTG	TGGTGT	TTTCT	CTGGGCTCACTGTTTC	960	
Db	901	ATTTG	TCCAGACT	TCAGGG	AAAGATG	GGTATTG	TGGTGT	TTTCT	CTGGGCTCACTGTTTC	960	
Qy	961	AAATG	CTTAC	AGAAAG	GGCTTAAT	CATTC	GCTCAG	CCCTTG	CCAGATCCCACAGA	1020	
Db	961	AAATG	CTTAC	AGAAAG	GGCTTAAT	CATTC	GCTCAG	CCCTTG	CCAGATCCCACAGA	1020	
Qy	1021	AGGTG	TTATG	SGAGTAC	AAAGAAAA	AAACCAT	CCACATT	AGGAG	CAATAC	TACGCGCTG	1080
Db	1021	AGGTG	TTATG	SGAGTAC	AAAGAAAA	AAACCAT	CCACATT	AGGAG	CAATAC	TACGCGCTG	1080
Qy	1081	ATGAT	TGGAT	TACCC	AGAAATG	ATCTT	CTGGT	CATCC	CAAGCAAG	CTTTATCACTC	1140
Db	1081	ATGAT	TGGAT	TACCC	AGAAATG	ATCTT	CTGGT	CATCC	CAAGCAAG	CTTTATCACTC	1140
Qy	1141	ATGGT	GGAATGA	TGGAT	CTATG	AGCTATT	TACCAT	GGGTCC	CTATCGT	GGGAGTTC	1200
Db	1141	ATGGT	GGAATGA	TGGAT	CTATG	AGCTATT	TACCAT	GGGTCC	CTATCGT	GGGAGTTC	1200
Qy	1201	CCATAT	TTG	TGATCAG	CTTGATA	AACATAG	CTCACAT	GAA	GGCCAAAGG	CAGAGCTGTAG	1260
Db	1201	CCATAT	TTG	TGATCAG	CTTGATA	AACATAG	CTCACAT	GAA	GGCCAAAGG	CAGAGCTGTAG	1260
Qy	1261	AAATA	AACTTCA	AAACTAT	GCAAG	GGAAGATTT	ACTG	AGG	GGCTTTC	GAGAAACAGTCAATTA	1320
Db	1261	AAATA	AACTTCA	AAACTAT	GCAAG	GGAAGATTT	ACTG	AGG	GGCTTTC	GAGAAACAGTCAATTA	1320
Qy	1321	CCGAT	TCTCTT	ATAAG	AGAATG	CTATG	AGATTAT	TCA	GAATTC	CACATGATCAACCTG	1380
Db	1321	CCGAT	TCTCTT	ATAAG	AGAATG	CTATG	AGATTAT	TCA	GAATTC	CACATGATCAACCTG	1380
Qy	1381	TAAAG	CCCCCTAG	ATCG	AGCAGTCTT	CTGG	ATCG	AGTTTGT	CATG	CGCCCCACAAAGGAGCCA	1440
Db	1381	TAAAG	CCCCCTAG	ATCG	AGCAGTCTT	CTGG	ATCG	AGTTTGT	CATG	CGCCCCACAAAGGAGCCA	1440

RESULT 4

US-10-176-758-521

03 10 176 758 521
: sequence 521. Application us/10176758

: Publication No. US20030008353A1

; PUBLICATION NO. US20
: GENERAL INFORMATION:

APPLICANT: Baker Kevin D

APPLICANT: BAKER, KEVIN

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc
APPLICANT: Coddard Andrew

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBR

; TITLE OF INVENTION: ACIDS E

; FILE REFERENCE: P3430R1C104

; CURRENT APPLICATION NUMBER: US/1

; CURRENT FILING DATE: 2002-06-21

; Prior Application removed

; NUMBER OF SEQ

; SEQ ID NO 521

; LENGTH: 2974

; TYPE: DNA

; ORGANISM: Homo

Query Match
100.0%: Score 1636; DB 9; Length 2974;

Query match	100.00%	100.00%
Best Local similarity	100.00%	100.00%
Best Local similarity	100.00%	100.00%

BEST LOCAL SIMILARITY 100.00%, FREQ: NO: 0,
 MATCHES 1636: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

1 GATCAGTGTGTGAGGGGAAC TGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTTCTGC 60

QY I GAICAGIGIGIGAGGGGACIACCATCAIAGAGGICICAGCAGCIIIGGIAIICIAGC
QY |||||

[illegible]

QY TCCAGCAGCTCTTCGTCGGTGTGGATCTGTGGGAAGTCCCTGGTGGCCCCCTGG 120

100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

QY 121 ACATGAGCCATTGGCTTAATGTCAAGGTCAATCTAGAAGAGCTCATAGTGAGAGGCCATG 180

181 AGGTAAACAGTATTGACTCACTCAAGCCTTCGTTAATTGACTACAGGAGCCTTCTGCAT 240 QY

QY 241 TGAAATTTGAGGTGTCATATGCCACAGGACAGAAATGAAATATTGTTG 300

0

301 ACCTAGCTCTGAATGTCTTGGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAAATG 360


```
QY 361 ATTTTGTGTTGAATGAAGGAACTTTAAATATGATGTGTGAGAGCTTTTATCTACAATC 420
Db 361 ATTTTGTGTTGAATGAAGGAACTTTAAATATGATGTGTGAGAGCTTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCTCTGA 480
Db 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCTCTGA 480
QY 481 TTCCCTGTGAGAGCTGATGGCTGAGTTGCTGTCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGAGAGCTGATGGCTGAGTTGCTGTCAGTCCCTTTTGTGCTCACACTTAGAA 540
QY 541 TTTCTGTAGGAGGCAATATGGAGCGAAGCTGTGGGAACCTTCCAGCTCCACTTTTCTATG 600
Db 541 TTTCTGTAGGAGGCAATATGGAGCGAAGCTGTGGGAACCTTCCAGCTCCACTTTTCTATG 600
QY 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAT 660
Db 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAT 660
QY 661 CAATGCTTTCAGTTTGTCTCCACTCTCGGATTCAGGATTCAGACTATCAATTTTGGGAAG 720
Db 661 CAATGCTTTCAGTTTGTCTCCACTCTCGGATTCAGGATTCAGACTATCAATTTTGGGAAG 720
QY 721 AGTTTATAGTAAGGCAATAGGAAGGCCCACTACATTTATGTGAGACTGTGGGAAGCTG 780
Db 721 AGTTTATAGTAAGGCAATAGGAAGGCCCACTACATTTATGTGAGACTGTGGGAAGCTG 780
QY 781 AGATATGGCTTAATACGAACATATTTGGATTTTGAATTTCTCAACCATACCAACCTAACT 840
Db 781 AGATATGGCTTAATACGAACATATTTGGATTTTGAATTTCTCAACCATACCAACCTAACT 840
QY 841 TTGAGTTTGTGGAGGATTCAGCTGTAAACCTGCCAAGCTTTGCCATAGGAAATGGAAA 900
Db 841 TTGAGTTTGTGGAGGATTCAGCTGTAAACCTGCCAAGCTTTGCCATAGGAAATGGAAA 900
QY 901 ATTTTGTCCAGAGTTCCAGGGAAGATGGTATTTGTGTTTCTCTGGGGTCACTGTTTC 960
Db 901 ATTTTGTCCAGAGTTCCAGGGAAGATGGTATTTGTGTTTCTCTGGGGTCACTGTTTC 960
QY 961 AAATGCTTACAGAAAGAAAGCTTAATCATTTGCTTACGCCCTTGCCAGATCCACAGA 1020
Db 961 AAATGCTTACAGAAAGAAAGCTTAATCATTTGCTTACGCCCTTGCCAGATCCACAGA 1020
QY 1021 AGGTGTTATGGAGGTACAAAGGAAACCAATCCACATTTAGAGCCATACCTCGCTGT 1080
Db 1021 AGGTGTTATGGAGGTACAAAGGAAACCAATCCACATTTAGAGCCATACCTCGCTGT 1080
QY 1081 ATGATTTGGATACCCAGATGATCTTTGGTTCATCCCAAAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTTGGATACCCAGATGATCTTTGGTTCATCCCAAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGGTGGAAATGAATGGGATCTATGAAGCTATTATACCATGGGTCCCTATGTGGAGTTC 1200
Db 1141 ATGGTGGATGAATGGGATCTATGAAGCTATTATACCATGGGTCCCTATGTGGAGTTC 1200
QY 1201 CCATATTTGGTGATCAGCTTGATAACATAGCTCAGATGAAGGCCAAGGAGCAGTGTAG 1260
Db 1201 CCATATTTGGTGATCAGCTTGATAACATAGCTCAGATGAAGGCCAAGGAGCAGTGTAG 1260
QY 1261 AAATAAATTCAAAACATGACAAAGGAGTATTTACTGAGGCTTTGAGAACAGTCAATTA 1320
Db 1261 AAATAAATTCAAAACATGACAAAGGAGTATTTACTGAGGCTTTGAGAACAGTCAATTA 1320
QY 1321 CCGATTTCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Db 1321 CCGATTTCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTAGATCGAGAGCTTCTTGGATCGAGTTTGTGATCGAGTTTGTGATCGAGGCA 1440
Db 1381 TAAAGCCCTAGATCGAGAGCTTCTTGGATCGAGTTTGTGATCGAGTTTGTGATCGAGGCA 1440
```

RESULT 5

```
US-10-175-737-521
; Sequence 521, Application US/10175737
; Publication NO. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deshoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; PRIOR FILING DATE: 2002-08-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-737-521
```

```
Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTGCG 60
Db 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTGCG 60
QY 61 TCCTGCAGCTCTCTGTGTTGGCTGTGGATTCTCTGGGAAAGTCCTGGTGTGGCCCTGTG 120
Db 61 TCCTGCAGCTCTCTGTGTTGGCTGTGGATTCTCTGGGAAAGTCCTGGTGTGGCCCTGTG 120
QY 121 ACATGAGCCATTTGACCTCAACTCAAGGCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTTGACCTCAACTCAAGGCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAACAGTATTGACCTCAACTCAAGGCTTCTAGAGAGCTCATAGTGAGAGGCCATG 240
Db 181 AGGTAACAGTATTGACCTCAACTCAAGGCTTCTAGAGAGCTCATAGTGAGAGGCCATG 240
QY 241 TGAATTTGAGGTGGTCCATATATGCGACGACAGACAGAGAAATGAAATATTGTTG 300
Db 241 TGAATTTGAGGTGGTCCATATATGCGACGACAGACAGAGAAATGAAATATTGTTG 300
QY 301 ACCTAGCTCGAATGCTTTCGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATG 360
Db 301 ACCTAGCTCGAATGCTTTCGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATG 360
```

Qy	361	ATTTTGTGTTGAAATTAAGAGGAACCTTTAAAAATGATGTGTGAGAGCTTTATCTTACAATC	420
Db			
Qy	361	ATTTTGTGTTGAAATTAAGAGGAACCTTTAAAAATGATGTGTGAGAGCTTTATCTTACAATC	420
Db			
Qy	421	AGAGCCTTATGAAGAAGCTACAGGAACCAACCTACGATGTAATGCTTTATAGACCCTGTGA	480
Db			
Qy	421	AGAGCCTTATGAAGAAGCTACAGGAACCAACCTACGATGTAATGCTTTATAGACCCTGTGA	480
Db			
Qy	481	TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA	540
Db			
Qy	481	TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA	540
Db			
Qy	541	TTTCTGTAGGAGGCAATATGGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCTATG	600
Db			
Qy	541	TTTCTGTAGGAGGCAATATGGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCTATG	600
Db			
Qy	601	TACCTGTGCCTATGACAGGACTTACACACAGAATGACCTTCTCTGGAAGAGTAAAAAATT	660
Db			
Qy	601	TACCTGTGCCTATGACAGGACTTACACACAGAATGACCTTCTCTGGAAGAGTAAAAAATT	660
Db			
Qy	661	CAATGCTTTTCAGTTTTTCCACTTCTCGGATTCAGGATTACGACTATCATTTTGGGAAG	720
Db			
Qy	661	CAATGCTTTTCAGTTTTTCCACTTCTCGGATTCAGGATTACGACTATCATTTTGGGAAG	720
Db			
Qy	721	AGTTTTATAGTAAGGCAATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAAGCTG	780
Db			
Qy	721	AGTTTTATAGTAAGGCAATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAAGCTG	780
Db			
Qy	781	AGATATGGCTAAATACGAACATATTTGGGATTTTGAATTTCTCTCAACCATACCACTTAAC	840
Db			
Qy	781	AGATATGGCTAAATACGAACATATTTGGGATTTTGAATTTCTCTCAACCATACCACTTAAC	840
Db			
Qy	841	TTGAGTTTGTGGAGGATTGACACTGTAAACCTGCGAAAGCTTTGCCCTTAAGGAAATGAAA	900
Db			
Qy	841	TTGAGTTTGTGGAGGATTGACACTGTAAACCTGCGAAAGCTTTGCCCTTAAGGAAATGAAA	900
Db			
Qy	901	ATTTTGTCCAGAGTTTCAGGGGAAGATGGTATTTGTGGTGTTTCTCTCGGGTCACTGTTTC	960
Db			
Qy	901	ATTTTGTCCAGAGTTTCAGGGGAAGATGGTATTTGTGGTGTTTCTCTCGGGTCACTGTTTC	960
Db			
Qy	961	AAAATGTTTACAGAAGAAAAGCTAAATATCATTTGCTTCAGCCCTTGCAGATCCACAGA	1020
Db			
Qy	961	AAAATGTTTACAGAAGAAAAGCTAAATATCATTTGCTTCAGCCCTTGCAGATCCACAGA	1020
Db			
Qy	1021	AGGTGTTTATGGAGTACAAAAGAAAAAACCATCCACATTAGGAGCCAACTACCTCGCTGT	1080
Db			
Qy	1021	AGGTGTTTATGGAGTACAAAAGAAAAAACCATCCACATTAGGAGCCAACTACCTCGCTGT	1080
Db			
Qy	1081	ATGATTGGATACCCCAAGTGAATCTTGGTCAATCCCAACCAAGAGCTTTTATCACTC	1140
Db			
Qy	1081	ATGATTGGATACCCCAAGTGAATCTTGGTCAATCCCAACCAAGAGCTTTTATCACTC	1140
Db			
Qy	1141	ATGGTGAATGAATGGGATCTATGAAGCTATTTTACCATGGGTCCTTATGGTGGAGTTC	1200
Db			
Qy	1141	ATGGTGAATGAATGGGATCTATGAAGCTATTTTACCATGGGTCCTTATGGTGGAGTTC	1200
Db			
Qy	1201	CCATATTTGTGTGATGAGCTTGATTAACATAGCTACATGAAGGCCAAAGGAGCAGCTGTAG	1260
Db			
Qy	1201	CCATATTTGTGTGATGAGCTTGATTAACATAGCTACATGAAGGCCAAAGGAGCAGCTGTAG	1260
Db			
Qy	1261	AAATAAACTTCAAACTATGACAGCGAAGATTTTACTTGGGGCTTTTGAGAACAGTCAATTA	1320
Db			
Qy	1261	AAATAAACTTCAAACTATGACAGCGAAGATTTTACTTGGGGCTTTTGAGAACAGTCAATTA	1320
Db			
Qy	1321	CCGATTCCTCTTATAAAGAGAATGCTATGAGATTTATCAAGAAATTCACATGATCAACCTG	1380
Db			
Qy	1321	CCGATTCCTCTTATAAAGAGAATGCTATGAGATTTATCAAGAAATTCACATGATCAACCTG	1380
Db			
Qy	1381	TAAAGCCCCCTAGATCGAGCAGCTTCTTGGATCGAGTTTGTATCGGCCCAAGAGAGCCA	1440
Db			
Qy	1381	TAAAGCCCCCTAGATCGAGCAGCTTCTTGGATCGAGTTTGTATCGGCCCAAGAGAGCCA	1440
Db			

RESULT 6

US-10-173-706-521

03 10 173 700 521
: Sequence 521, Application US/10173706

; Publication No. US20030022293A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey

APPLICANT: GODDARD, Audrey
APPLICANT: Godowski, Paul J.

APPLICANT: GURNEY, PAUL J.
APPLICANT: GURNEY, AUSTIN L.

APPLICANT: GURNEY, AUSTIN
APPLICANT: Pan, James

APPLICANT: PAUL, JAMES
APPLICANT: SMITH, VICTORIA

APPLICANT: SMITH, VICTORIA
APPLICANT: WATERABE COLIN K

APPLICANT: Watanabe, Collin K
APPLICANT: Wood William T

APPLICANT: WOOD, WILLIAM
APPLICANT: ZHANG ZONGJIN

APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRATED ENDOGENOUS THE GENE

DATE	DESCRIPTION	AMOUNT	TITLE OF INVENTION
1900	1000	1000	1000
1901	1000	1000	1000
1902	1000	1000	1000
1903	1000	1000	1000
1904	1000	1000	1000
1905	1000	1000	1000
1906	1000	1000	1000
1907	1000	1000	1000
1908	1000	1000	1000
1909	1000	1000	1000
1910	1000	1000	1000
1911	1000	1000	1000
1912	1000	1000	1000
1913	1000	1000	1000
1914	1000	1000	1000
1915	1000	1000	1000
1916	1000	1000	1000
1917	1000	1000	1000
1918	1000	1000	1000
1919	1000	1000	1000
1920	1000	1000	1000
1921	1000	1000	1000
1922	1000	1000	1000
1923	1000	1000	1000
1924	1000	1000	1000
1925	1000	1000	1000
1926	1000	1000	1000
1927	1000	1000	1000
1928	1000	1000	1000
1929	1000	1000	1000
1930	1000	1000	1000
1931	1000	1000	1000
1932	1000	1000	1000
1933	1000	1000	1000
1934	1000	1000	1000
1935	1000	1000	1000
1936	1000	1000	1000
1937	1000	1000	1000
1938	1000	1000	1000
1939	1000	1000	1000
1940	1000	1000	1000
1941	1000	1000	1000
1942	1000	1000	1000
1943	1000	1000	1000
1944	1000	1000	1000
1945	1000	1000	1000
1946	1000	1000	1000
1947	1000	1000	1000
1948	1000	1000	1000
1949	1000	1000	1000
1950	1000	1000	1000
1951	1000	1000	1000
1952	1000	1000	1000
1953	1000	1000	1000
1954	1000	1000	1000
1955	1000	1000	1000
1956	1000	1000	1000
1957	1000	1000	1000
1958	1000	1000	1000
1959	1000	1000	1000
1960	1000	1000	1000
1961	1000	1000	1000
1962	1000	1000	1000
1963	1000	1000	1000
1964	1000	1000	1000
1965	1000	1000	1000
1966	1000	1000	1000
1967	1000	1000	1000
1968	1000	1000	1000
1969	1000	1000	1000
1970	1000	1000	1000
1971	1000	1000	1000
1972	1000	1000	1000
1973	1000	1000	1000
1974	1000	1000	1000
1975	1000	1000	1000
1976	1000	1000	1000
1977	1000	1000	1000
1978	1000	1000	1000
1979	1000	1000	1000
1980	1000	1000	1000
1981	1000	1000	1000
1982	1000	1000	1000
1983	1000	1000	1000
1984	1000	1000	1000
1985	1000	1000	1000

```

; FILE REFERENCE: P3430R1C7
;

```

; CURRENT APPLICATION NUMBER: US/1

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed

; NUMBER OF SEQ

; SEQ ID NO 521

; LENGTH: 2

; TYPE: DNA

; ORGANISM: Homo

Query Match 100.0%: Score 1636: DB 9: Length 2974:

Query Match	100.0%,	Score	1030,
Best Local Similarity:	100.0%	pred	No. 0:

BEST LOCAL SIMILARITY 100.0%; PRED. NO: 0;
Matches 1636: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

1 GATCAGTGTGTGAGGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTGGTATTCTGC 60

1

120

61 TCC TGC AGC T C T C T G T G T G G C T G T G G A T C T G T G G G A A G T C C T G G T G T G G C C T G T G 120

Qy 121 ACATGAGCCATTGGCTTAATGTCAAGGTCATTCTAGAAGAGCTCATAGTGAGAGGCCATG 180

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

181 AGGTAACAGTATTGACTCACTCAAAGCCTTCGTTAATTGACTACAGGAGCCTTCTGCAT 240

DD TGGTAAACAGTATGGACATCATTCAATGAGCTTCGTATATATGGACATACAGGAAAGCCCTTCCTCCAT 249

241 TCAATTTCCCTCCATGCCACACGACACAAATGAAATATTGTTC 300

Db 241 TGAAATTGAGGTGGTCCATATGCCACAGGACAGAAATGAAAATATTTGTTG 300

```
QY 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATAATGATGTGTGAGAGCTTTATCTACAATC 420
DB 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATAATGATGTGTGAGAGCTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAGCTACAGGAACCACTACGATGTATGCTTTATAGACCCCTGTGA 480
DB 421 AGACGCTTATGAAGAGCTACAGGAACCACTACGATGTATGCTTTATAGACCCCTGTGA 480
QY 481 TTCCTGTGGAGACCTGATGGCTGAGTTGCTTGGAGTCCCTTTTGTGCTCAGCTTAGAA 540
DB 481 TTCCTGTGGAGACCTGATGGCTGAGTTGCTTGGAGTCCCTTTTGTGCTCAGCTTAGAA 540
QY 541 TTTCTGTAGGAGCAATATGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTCTATG 600
DB 541 TTTCTGTAGGAGCAATATGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTCTATG 600
QY 601 TACCTGTGCTATGACAGACTAACAGAGCAATGACCTTTCTGGGAAAGTAAATAAT 660
DB 601 TACCTGTGCTATGACAGACTAACAGAGCAATGACCTTTCTGGGAAAGTAAATAAT 660
QY 661 CAATGCTTTCAGTTTGTTCACCTTCTGATTCAGGATACGACTATCATTTTGGGAG 720
DB 661 CAATGCTTTCAGTTTGTTCACCTTCTGATTCAGGATACGACTATCATTTTGGGAG 720
QY 721 AGTTTATAGTAAGGCAATAGGAAGCCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
DB 721 AGTTTATAGTAAGGCAATAGGAAGCCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
QY 781 AGATATGCTAATACGAACATATGGGATTTTGAATTTCCCTCAACCATACCAACCTAAT 840
DB 781 AGATATGCTAATACGAACATATGGGATTTTGAATTTCCCTCAACCATACCAACCTAAT 840
QY 841 TTGAGTTTGTGGAGGATTCAGCTTAACCTGCCAAGCTTTGCTTAAGGAATGGAA 900
DB 841 TTGAGTTTGTGGAGGATTCAGCTTAACCTGCCAAGCTTTGCTTAAGGAATGGAA 900
QY 901 ATTTTGTCCAGAGTTTCAGGGGAAGATGTTATGTGTTGTTCTCTGGGGTTCACGTTC 960
DB 901 ATTTTGTCCAGAGTTTCAGGGGAAGATGTTATGTGTTGTTCTCTGGGGTTCACGTTC 960
QY 961 AAATGTTACAGAAAGAGCTTAATATCATTTGCTTACGCCCTTCCCGAGATCCACAGA 1020
DB 961 AAATGTTACAGAAAGAGCTTAATATCATTTGCTTACGCCCTTCCCGAGATCCACAGA 1020
QY 1021 AGGCTCTTATGGAGGTACAAAGGAAAAAACCACATACGATAGGAGCCATCTCGGCTGT 1080
DB 1021 AGGCTCTTATGGAGGTACAAAGGAAAAAACCACATACGATAGGAGCCATCTCGGCTGT 1080
QY 1081 ATGATTTGGATACCCAGAAATGATCTTTTGTGTCATCCCAAAACCAAGCTTTTATCACTC 1140
DB 1081 ATGATTTGGATACCCAGAAATGATCTTTTGTGTCATCCCAAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGGTGGAAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGTGGAGTTC 1200
DB 1141 ATGGTGGAAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGTGGAGTTC 1200
QY 1201 CCATATTTGGTGATCAGCTTGATACATAGCTCAGATGAAGCCCAAGGAGCAGCTGTAG 1260
DB 1201 CCATATTTGGTGATCAGCTTGATACATAGCTCAGATGAAGCCCAAGGAGCAGCTGTAG 1260
QY 1261 AAATAAAGCTTCAAAACTATGACAAGCGAAGATTTTACTGAGGGCTTTGAGACAGTCATTA 1320
DB 1261 AAATAAAGCTTCAAAACTATGACAAGCGAAGATTTTACTGAGGGCTTTGAGACAGTCATTA 1320
QY 1321 CCGATTCCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
DB 1321 CCGATTCCTCTTATAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGATCGAGTTTGTCTATGCGCCCAAGAGGCA 1440
DB 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGATCGAGTTTGTCTATGCGCCCAAGAGGCA 1440
```

```
QY 1441 AGCACCCTGCGATCAGCTGCCATGACCTCACCTGGTTCCAGCAGCTACTCTATAGATGTGA 1500
DB 1441 AGCACCCTGCGATCAGCTGCCATGACCTCACCTGGTTCCAGCAGCTACTCTATAGATGTGA 1500
QY 1501 TTGGTTTCTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
DB 1501 TTGGTTTCTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
QY 1561 TTTCTGTCAAAAATTTTAAATAAAGTAAAGAGATAGAAAAGGGAATAGATCTTTCCAA 1620
DB 1561 TTTCTGTCAAAAATTTTAAATAAAGTAAAGAGATAGAAAAGGGAATAGATCTTTCCAA 1620
QY 1621 ATTCAGAAAGACCTG 1636
DB 1621 ATTCAGAAAGACCTG 1636

RESULT 7
US-10-175-738-521
; Sequence 521, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTTGC 60
DB 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTCTTGC .60
QY 61 TCCTCAGCTCTTCTGTGTGGCTGTGGATTCTGTGGGAAAGTCCCTGGTGTGGCCCTGTG 120
DB 61 TCCTCAGCTCTTCTGTGTGGCTGTGGATTCTGTGGGAAAGTCCCTGGTGTGGCCCTGTG 120
QY 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
DB 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAACAGTATTGACTCAGCTCAAGGCTTCGTTAATTGACTACAGAGCCCTTCTGCAT 240
DB 181 AGGTAACAGTATTGACTCAGCTCAAGGCTTCGTTAATTGACTACAGAGCCCTTCTGCAT 240
QY 241 TGAATTTTGGGTGGTCCATATGCCACAGGACAGAACAGAAAATGAAATATTTTGTG 300
DB 241 TGAATTTTGGGTGGTCCATATGCCACAGGACAGAACAGAAAATGAAATATTTTGTG 300
QY 301 ACCTAGCTCTGAATCTCTTGGCAGGCTTATCAACCTGGCAATACAGTTATAAATTTAAATG 360
DB 301 ACCTAGCTCTGAATCTCTTGGCAGGCTTATCAACCTGGCAATACAGTTATAAATTTAAATG 360
```

Qy	361	ATTTTGTGTTGAAATAAGAGAACTTTAAAAATGATGCTGAGAGCTTTATCTACAATC	420
Db			
Qy	361	ATTTTGTGTTGAAATAAGAGAACTTTAAAAATGATGCTGAGAGCTTTATCTACAATC	420
Db			
Qy	421	AGAGCTTATGAGAAGCTACAGAAACCAACTACGATGAATGCTTTATAGACCTGTGTA	480
Db			
Qy	421	AGAGCTTATGAGAAGCTACAGAAACCAACTACGATGAATGCTTTATAGACCTGTGTA	480
Db			
Qy	481	TTCCCTGTGAGACCTGATGCTGAGTTGCTTCAGTCCCTTTTGCTCACACTTAGAA	540
Db			
Qy	481	TTCCCTGTGAGACCTGATGCTGAGTTGCTTCAGTCCCTTTTGCTCACACTTAGAA	540
Db			
Qy	541	TTTCTGTAGGAGCAATATGAGCGAAGCTGTGGGAACTTCCAGTCCCACTTTCCTATG	600
Db			
Qy	541	TTTCTGTAGGAGCAATATGAGCGAAGCTGTGGGAACTTCCAGTCCCACTTTCCTATG	600
Db			
Qy	601	TACTGTGCTATGACAGGACTAACAGACAGAATGACCTTTCTGGAAAGAGTAAAAAAT	660
Db			
Qy	601	TACTGTGCTATGACAGGACTAACAGACAGAATGACCTTTCTGGAAAGAGTAAAAAAT	660
Db			
Qy	661	CAATGCTTTTCAGTTTGTCCACTCTCGGATTCAGGATTCAGTACTATCATTTTGGGAAG	720
Db			
Qy	661	CAATGCTTTTCAGTTTGTCCACTCTCGGATTCAGGATTCAGTACTATCATTTTGGGAAG	720
Db			
Qy	721	AGTTTTATAGTAAAGCAATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAAGCTG	780
Db			
Qy	721	AGTTTTATAGTAAAGCAATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAAGCTG	780
Db			
Qy	781	AGATATGGCTTAATACGAACATATTTGGGATTTTGAATTTCTCAACCATCAACCTAACT	840
Db			
Qy	781	AGATATGGCTTAATACGAACATATTTGGGATTTTGAATTTCTCAACCATCAACCTAACT	840
Db			
Qy	841	TTGAGTTTGTGGAGGATTCGACTGTAAAGCTGCGGAAAGCTTTGCTGAGGAAATGGAAA	900
Db			
Qy	841	TTGAGTTTGTGGAGGATTCGACTGTAAAGCTGCGGAAAGCTTTGCTGAGGAAATGGAAA	900
Db			
Qy	901	ATTTTGTCCAGAGTTTCAGGGGAAGATGGTATTTGTGGTGTTCCTCTGGGCTCACTGTTTC	960
Db			
Qy	901	ATTTTGTCCAGAGTTTCAGGGGAAGATGGTATTTGTGGTGTTCCTCTGGGCTCACTGTTTC	960
Db			
Qy	961	AAAATGTTACAGAGAAAAGGCTTAATATCATTTGCTTACGCCCTTGCCAGATCCCAAGA	1020
Db			
Qy	961	AAAATGTTACAGAGAAAAGGCTTAATATCATTTGCTTACGCCCTTGCCAGATCCCAAGA	1020
Db			
Qy	1021	AGGTGTTATCGAGGTACAAAGGAAAAAACCATCCACATTAGAGGCCAATACTCGGCTGT	1080
Db			
Qy	1021	AGGTGTTATCGAGGTACAAAGGAAAAAACCATCCACATTAGAGGCCAATACTCGGCTGT	1080
Db			
Qy	1081	ATGATTTGGATACCCAGAAATGATCTTCTGGTTCATCCAAAACCAAAGCTTTTATCACTC	1140
Db			
Qy	1081	ATGATTTGGATACCCAGAAATGATCTTCTGGTTCATCCAAAACCAAAGCTTTTATCACTC	1140
Db			
Qy	1141	ATGGTGAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGGAGTTC	1200
Db			
Qy	1141	ATGGTGAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGGAGTTC	1200
Db			
Qy	1201	CCATATTTGGTGATCAGCTTGTAAACATGCTCAGATGAAGGCCAAAGGAGACGTGTAG	1260
Db			
Qy	1201	CCATATTTGGTGATCAGCTTGTAAACATGCTCAGATGAAGGCCAAAGGAGACGTGTAG	1260
Db			
Qy	1261	AAATAAACTTCAAACCTATCACAGGCGAAGATTTACTGAGGGCTTTTGAGAACAGTCAATTA	1320
Db			
Qy	1261	AAATAAACTTCAAACCTATCACAGGCGAAGATTTACTGAGGGCTTTTGAGAACAGTCAATTA	1320
Db			
Qy	1321	CCGATTCCTCTTATAAGAGAAATGCTATGAGATTTCAAGAATTCACCATGATCAACCTG	1380
Db			
Qy	1321	CCGATTCCTCTTATAAGAGAAATGCTATGAGATTTCAAGAATTCACCATGATCAACCTG	1380
Db			
Qy	1381	TAAAGCCCCCTAGATTCGAGCAGTCTTCTGGGATCGAGTTTGTATGCGGCCAAAGGAGCCA	1440
Db			
Qy	1381	TAAAGCCCCCTAGATTCGAGCAGTCTTCTGGGATCGAGTTTGTATGCGGCCAAAGGAGCCA	1440
Db			

RESULT 8

RESULTS
US-10-175-752-521

US-10-173-732-321
: Sequence 521, Application US/10175752.

Sequence 321, Application US/10030002295A1
Publication No. US20030022295A1

: PUBLICATION NO: 0320
: GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: CHEN, JIAN

APPLICANT: Desnoyers, Luc

APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: zhang, zemin

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: 0242001060

; FILE REFERENCE: P3430R1C60
: CURRENT ADDITION NUMBER:

```

; CURRENT APPLICATION NUMBER: US/10/175,752
: CURRENT FILING DATE: 2002-06-19

```

; CURRENT FILING DATE: 2002-06-19
: Prior Application removed - See

; Prior Application removed - See File Wrapper or Palm
 : NUMBER OF SEQ ID NOS: 612

```

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521

```

```

; SEQ ID NO 521
; LENGTH: 2974

```

```

; LENGTH: 2974
; TYPE: DNA

```

; LIFE: DNA
; ORGANISM:

US-10-175-752-521

4
4
2
3
2
2
4
2
4
2
2
2

Query Match 100.0%; Score 1636; DB 9; Length 2974;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GATCAGTCTGTGAGGGAAC	TGCCATCATGAGTCTGACAAGT	CAGCTTGTGTTATTTCTGC	60
Db	1	GATCAGTCTGTGAGGGAAC	TGCCATCATGAGTCTGACAAGT	CAGCTTGTGTTATTTCTGC	60
Qy	61	TCCTGCAGCTCTTCTGTGTTGGCTGTGGATTC	TCTGTGGGAAAGTCTCGTGTGGCCCTGTG	120	
Db	61	TCCTGCAGCTCTTCTGTGTTGGCTGTGGATTC	TCTGTGGGAAAGTCTCGTGTGGCCCTGTG	120	
Qy	121	ACATGAGCCATTTGGCTTTAATGTCAAGGTCA	TTCCTAGAGAGCTCATAGTGAGAGCCATG	180	
Db	121	ACATGAGCCATTTGGCTTTAATGTCAAGGTCA	TTCCTAGAGAGCTCATAGTGAGAGCCATG	180	
Qy	181	AGGTAACAGATTTGACACTCAAGGCTTCGTTAA	TTGACTACAGGAGGCTTCTGCAT	240	
Db	181	AGGTAACAGATTTGACACTCAAGGCTTCGTTAA	TTGACTACAGGAGGCTTCTGCAT	240	
Qy	241	TGAAATTTGAGTGCGTCCATATGCCAGGACACA	AGAGAAATCAAAATATTTGTTG	300	
Db	241	TGAAATTTGAGTGCGTCCATATGCCAGGACACA	AGAGAAATCAAAATATTTGTTG	300	
Qy	301	ACCTAGCTCTGAATGTCTTGCCAGGCTTATCA	AGCTGGCAATCAGTTATAAAATTAATG	360	
Db	301	ACCTAGCTCTGAATGTCTTGCCAGGCTTATCA	AGCTGGCAATCAGTTATAAAATTAATG	360	

QY 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATATGATGCTGAGAGCTTTTATCTACAATC 420
DB 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATATGATGCTGAGAGCTTTTATCTACAATC 420
QY 421 AGACGGCTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCCGTGA 480
DB 421 AGACGGCTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCCGTGA 480
QY 481 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTTCGAGTCCCTTTTGTGCTCACACTTAGAA 540
DB 481 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTTCGAGTCCCTTTTGTGCTCACACTTAGAA 540
QY 541 TTTCTGTAGGAGCAATATAGGAGCAAGCTGTGGAACTTCAGGCTCCACTTTTCTTATG 600
DB 541 TTTCTGTAGGAGCAATATAGGAGCAAGCTGTGGAACTTCAGGCTCCACTTTTCTTATG 600
QY 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAAGAGTAAATAAT 660
DB 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAAGAGTAAATAAT 660
QY 661 CAATGCTTTCAGTTTCTCCACTTCTGGATTGAGGATACGACTATCATTTTGGGAAG 720
DB 661 CAATGCTTTCAGTTTCTCCACTTCTGGATTGAGGATACGACTATCATTTTGGGAAG 720
QY 721 AGTTTATAGTAAAGCAATTAGGAAGCCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
DB 721 AGTTTATAGTAAAGCAATTAGGAAGCCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
QY 781 AGATATGGCTTAATACGAACATATGGAATTTGAAATTTCTCAACCATACCAACTAAT 840
DB 781 AGATATGGCTTAATACGAACATATGGAATTTGAAATTTCTCAACCATACCAACTAAT 840
QY 841 TTGAGTTTGTGGAGGATTCGACTGTAACTGCAAGCTTTCCTTAAGAAATGAAA 900
DB 841 TTGAGTTTGTGGAGGATTCGACTGTAACTGCAAGCTTTCCTTAAGAAATGAAA 900
QY 901 ATTTTGTCCAGAGTTACAGGGAAGATGATGTTGGTGTGTTTCTCTGGGTCAGTGTTC 960
DB 901 ATTTTGTCCAGAGTTACAGGGAAGATGATGTTGGTGTGTTTCTCTGGGTCAGTGTTC 960
QY 961 AAATGTTACAGAAGAAAGCTTAATATCATTTGCTCAGCCCTTGCCAGATCCACAGA 1020
DB 961 AAATGTTACAGAAGAAAGCTTAATATCATTTGCTCAGCCCTTGCCAGATCCACAGA 1020
QY 1021 AGGTGTTATGAGGTACAAAGGAAACCATCCATAGGAGCAATACTCGGCTGT 1080
DB 1021 AGGTGTTATGAGGTACAAAGGAAACCATCCATAGGAGCAATACTCGGCTGT 1080
QY 1081 ATGATTGGATACCCAGATGATCTTCTTGCTATCCCAAAACCAAGCTTTTATCACTC 1140
DB 1081 ATGATTGGATACCCAGATGATCTTCTTGCTATCCCAAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGTTGGAATGAATGGATCTATGAAGCTATTTACATGGGGTCCCTATGGTGGAGTTC 1200
DB 1141 ATGTTGGAATGAATGGATCTATGAAGCTATTTACATGGGGTCCCTATGGTGGAGTTC 1200
QY 1201 CCATATTTGTGATCAGCTTGATACATAGCTACATGAAGGCCAAAGGAGCACTGTAG 1260
DB 1201 CCATATTTGTGATCAGCTTGATACATAGCTACATGAAGGCCAAAGGAGCACTGTAG 1260
QY 1261 AAATAAATCTTCAAACTATGACAGGAAGATTTACTGAGGCTTTGAGACAGCTCATTA 1320
DB 1261 AAATAAATCTTCAAACTATGACAGGAAGATTTACTGAGGCTTTGAGACAGCTCATTA 1320
QY 1321 CCGATTCCTCTTATAAGAGAAATGCTATGAGATTATCAAGAAATTCACCATGATCAACCTG 1380
DB 1321 CCGATTCCTCTTATAAGAGAAATGCTATGAGATTATCAAGAAATTCACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTTAGATCGAGCAGCTCTTCTGGATCGAGTTTGTGTCATGGCCCAAAAGAGCCA 1440
DB 1381 TAAAGCCCTTAGATCGAGCAGCTCTTCTGGATCGAGTTTGTGTCATGGCCCAAAAGAGCCA 1440

RESULT 9

US-10-176-482-521
; Sequence 521, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-482-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGTCTGACAGTCAGCTTTGTTATTTCTGC 60
DB 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGTCTGACAGTCAGCTTTGTTATTTCTGC 60
QY 61 TCCTGACGCTTCTGTTGGCTGTGATTTCTGTGGAAAGTCTGTTGGTGGCCCTGTG 120
DB 61 TCCTGACGCTTCTGTTGGCTGTGATTTCTGTGGAAAGTCTGTTGGTGGCCCTGTG 120
QY 121 ACATGAGCCATTGGCTTAATGTCAGGTCATCTTAGAGAGCTCATAGTGAGAGGCCATG 180
DB 121 ACATGAGCCATTGGCTTAATGTCAGGTCATCTTAGAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAAACAGTATTCACCTACCTCAAGCCCTTCTTAATTTGCTACAGGAAGCTTCTGCAT 240
DB 181 AGGTAAACAGTATTCACCTACCTCAAGCCCTTCTTAATTTGCTACAGGAAGCTTCTGCAT 240
QY 241 TGAATTTGAGGTGGTCCATATGCCACAGGACAGACAGAAAGTAAATATTTGTTG 300
DB 241 TGAATTTGAGGTGGTCCATATGCCACAGGACAGACAGAAAGTAAATATTTGTTG 300
QY 301 ACCTAGCTCTGAATGCTTGTGCGAGGCTTATCAACCTGGCAATCAGTTATAAATTAATG 360
DB 301 ACCTAGCTCTGAATGCTTGTGCGAGGCTTATCAACCTGGCAATCAGTTATAAATTAATG 360

```
QY 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATATGATGTGTGAGAGCTTTATCTACAATC 420
Db 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATATGATGTGTGAGAGCTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTTATAGACCCGTGTA 480
Db 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTTATAGACCCGTGTA 480
QY 481 TTCCCTGTGGAGACCTGATGGCTGAGTGTCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGGAGACCTGATGGCTGAGTGTCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
QY 541 TTCTCTAGGAGCAATATGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCATG 600
Db 541 TTCTCTAGGAGCAATATGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCATG 600
QY 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAAT 660
Db 601 TACCTGTGCTATGACAGGACTAACAGACAGATGACCTTTCTGGAAGAGTAAAAAAT 660
QY 661 CAATGCTTTCAGTTTGTCCACTTCTGGATTCAGGATTCAGGACTATCATTTTGGGAAG 720
Db 661 CAATGCTTTCAGTTTGTCCACTTCTGGATTCAGGATTCAGGACTATCATTTTGGGAAG 720
QY 721 AGTTTATATAGAGCATTAGGAAGCCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
Db 721 AGTTTATATAGAGCATTAGGAAGCCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
QY 781 AGATATGGCTATACGAACATATGGGATTTTGAATTTCTCAACCATACCAACCTTAAT 840
Db 781 AGATATGGCTATACGAACATATGGGATTTTGAATTTCTCAACCATACCAACCTTAAT 840
QY 841 TTGAGTTTGTGGAGGATTCAGCTGTAAACCTGCCAAAGCTTTGCCCTAAGGAATGAAA 900
Db 841 TTGAGTTTGTGGAGGATTCAGCTGTAAACCTGCCAAAGCTTTGCCCTAAGGAATGAAA 900
QY 901 ATTTTGTCCAGAGTTTCCAGGGAAGATGGTATTTGTGTGTTTCTCTGGGCTCACTTTTC 960
Db 901 ATTTTGTCCAGAGTTTCCAGGGAAGATGGTATTTGTGTGTTTCTCTGGGCTCACTTTTC 960
QY 961 AATATCTTACAGAGAAGAGCTTAATATCATTTGCTTCAGCCCTTGCCCAAGTCCACAGA 1020
Db 961 AATATCTTACAGAGAAGAGCTTAATATCATTTGCTTCAGCCCTTGCCCAAGTCCACAGA 1020
QY 1021 AGGTGTTATGGAGGTACAAAGGAAAAAACCACTCCACATTAGGAGCAATACCTCGCTGT 1080
Db 1021 AGGTGTTATGGAGGTACAAAGGAAAAAACCACTCCACATTAGGAGCAATACCTCGCTGT 1080
QY 1081 ATGATTTGGATACCCCAAGATGATCTTCTTGGTCAATCCCAAAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTTGGATACCCCAAGATGATCTTCTTGGTCAATCCCAAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGGTGGATGAATGGGATCTATGAAGCTATTTACCATGGGTCCCTATGTTGGGAGTTC 1200
Db 1141 ATGGTGGATGAATGGGATCTATGAAGCTATTTACCATGGGTCCCTATGTTGGGAGTTC 1200
QY 1201 CCATATTTGGTATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGAGCAGCTGTAG 1260
Db 1201 CCATATTTGGTATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGAGCAGCTGTAG 1260
QY 1261 AATAAATCTCAAACTATGACAAAGCAAGATTTACTGAGGGCTTTTGAGAACAGTCAATTA 1320
Db 1261 AATAAATCTCAAACTATGACAAAGCAAGATTTACTGAGGGCTTTTGAGAACAGTCAATTA 1320
QY 1321 CGATTTCTCTTATAAAGAGATGCTATGAGATTTATCAAGATTTACCATGATCAACCTG 1380
Db 1321 CGATTTCTCTTATAAAGAGATGCTATGAGATTTATCAAGATTTACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTTAGATCAGGAGCTTCTTGGATCGAGTTTGTCTGATCGAGTTTGTCTGCGCCACAAAGAGGCA 1440
Db 1381 TAAAGCCCTTAGATCAGGAGCTTCTTGGATCGAGTTTGTCTGATCGAGTTTGTCTGCGCCACAAAGAGGCA 1440
```

```
QY 1441 AGCACCTGCGATCAGCTGCCATGACCTCACCTGTTCCAGCACTACTCTATAGATGTA 1500
Db 1441 AGCACCTGCGATCAGCTGCCATGACCTCACCTGTTCCAGCACTACTCTATAGATGTA 1500
QY 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
Db 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATTTCTTTCACAAAATGTTTTTAT 1560
QY 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1620
Db 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1620
QY 1621 ATTCAAGAAGACCTG 1636
Db 1621 ATTCAAGAAGACCTG 1636
```

RESULT 10

```
US-10-176-757-521
; Sequence 521, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-521
```

```
Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTTCTGC 60
Db 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTTCTGC 60
QY 61 TCCTGAGCTCTTCTGTTGGCTGTGGATTTCTGTGGGAAAGTCCTGGTGGCCCTGTG 120
Db 61 TCCTGAGCTCTTCTGTTGGCTGTGGATTTCTGTGGGAAAGTCCTGGTGGCCCTGTG 120
QY 121 ACATGAGCCATTGGCTTAATGTCAAGTCTATTTCTAGAAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTGGCTTAATGTCAAGTCTATTTCTAGAAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAACAGTATTGACTCACTCAAAAGCTTTGTTAAATGACTACAGAAAGCCCTTCTGCAT 240
Db 181 AGGTAACAGTATTGACTCACTCAAAAGCTTTGTTAAATGACTACAGAAAGCCCTTCTGCAT 240
QY 241 TGAATTTGAGTGGTCCATATGCCACAGGACAGACAGAAAATGAATATTTGTTG 300
Db 241 TGAATTTGAGTGGTCCATATGCCACAGGACAGACAGAAAATGAATATTTGTTG 300
QY 301 ACCTAGCTCTGAATGTCTTGGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATG 360
Db 301 ACCTAGCTCTGAATGTCTTGGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATG 360
```



```
QY 361 ATTTTCTTCTGTAATAGAGGAACCTTTAAATATGATGTGTGAGACCTTTATCTACAATC 420
Db 361 ATTTTCTTCTGTAATAGAGGAACCTTTAAATATGATGTGTGAGACCTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAAAGCTACAGGAACCAACTACGATGATGCTTTATAGACCCCTGTGA 480
Db 421 AGACGCTTATGAAGAAAGCTACAGGAACCAACTACGATGATGCTTTATAGACCCCTGTGA 480
QY 481 TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
QY 541 TTCTCTGTAGAGGCAATATGGAGCGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 600
Db 541 TTCTCTGTAGAGGCAATATGGAGCGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 600
QY 601 TACCTGTGCTATGACAGGACTAACAGACAGAAATGACCTTCTGGAAGAGTAAAAAAT 660
Db 601 TACCTGTGCTATGACAGGACTAACAGACAGAAATGACCTTCTGGAAGAGTAAAAAAT 660
QY 661 CAATGCTTTCAGTTTGTCCACTTCTGGATTCCAGGATACGACTATCATTTTGGGAAG 720
Db 661 CAATGCTTTCAGTTTGTCCACTTCTGGATTCCAGGATACGACTATCATTTTGGGAAG 720
QY 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
Db 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTATGTGAGACTGTGGGAAAGCTG 780
QY 781 AGATATGGCTTAATACGAACATATGGGATTTTGAATTTTCTCAACCATACCAACCTAACT 840
Db 781 AGATATGGCTTAATACGAACATATGGGATTTTGAATTTTCTCAACCATACCAACCTAACT 840
QY 841 TTGAGTTTGTGAGGATTCACCTGTAAACCTCCCAAGCTTTGCTCAAGGAATGAAA 900
Db 841 TTGAGTTTGTGAGGATTCACCTGTAAACCTCCCAAGCTTTGCTCAAGGAATGAAA 900
QY 901 ATTTTGTCCAGACTTCAGGGAAGATGTTGTTGTTCTCTGCGGCTCACTGTTTC 960
Db 901 ATTTTGTCCAGACTTCAGGGAAGATGTTGTTGTTCTCTGCGGCTCACTGTTTC 960
QY 961 AAAATGTTACAGAAAGAGCTTAATATCATCTTCAGCCCTTGCCAGATCCACAGA 1020
Db 961 AAAATGTTACAGAAAGAGCTTAATATCATCTTCAGCCCTTGCCAGATCCACAGA 1020
QY 1021 AGTGTTATGAGGATACAAAGGAAAGAAACCATTCACATTAGGAGCCATACCTCGCTGT 1080
Db 1021 AGTGTTATGAGGATACAAAGGAAAGAAACCATTCACATTAGGAGCCATACCTCGCTGT 1080
QY 1081 ATGATTGGATACCCAGAAATGATCTTCTGCTCATCCAAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTGGATACCCAGAAATGATCTTCTGCTCATCCAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGGTGAATGAATGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGAGTTC 1200
Db 1141 ATGGTGAATGAATGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGAGTTC 1200
QY 1201 CCATATTTGCTGATCAGCTTGTAAACATAGCTCACATGAAGGCCAAGAGGAGCTGTAG 1260
Db 1201 CCATATTTGCTGATCAGCTTGTAAACATAGCTCACATGAAGGCCAAGAGGAGCTGTAG 1260
QY 1261 AAATAAACTTCAAACTATGACAAAGCAAGATTTACTGAGGGCTTTTGAGAACAGTCATTA 1320
Db 1261 AAATAAACTTCAAACTATGACAAAGCAAGATTTACTGAGGGCTTTTGAGAACAGTCATTA 1320
QY 1321 CCGATTCCCTTTTAAAGAGAATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTG 1380
Db 1321 CCGATTCCCTTTTAAAGAGAATGCTATGAGATTTATCAAGAATTCACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTATCGCCGCCCAAGAGGCCA 1440
Db 1381 TAAAGCCCTTAGATCGAGCAGTCTTCTGGATCGAGTTTGTATCGCCGCCCAAGAGGCCA 1440

QY 1441 AGCACCCTGCGATCAGCTGCCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTA 1500
Db 1441 AGCACCCTGCGATCAGCTGCCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTA 1500
QY 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTTAT 1560
Db 1501 TTGGGTTCTGCTGACCTGTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTTAT 1560
QY 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1620
Db 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTCCAA 1620
QY 1621 ATTCAAGAAAGACCTG 1636
Db 1621 ATTCAAGAAAGACCTG 1636

RESULT 11
US-10-176-913-521
: Sequence 521, Application US/10176913
: Publication No. US20030022298A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C66
: CURRENT FILING DATE: 2002-06-20
: Prior Application removed - See file Wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 521
: LENGTH: 2974
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-913-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGTGTGTGAGGGAACCTCCATCATGAGGCTGACAAAGTCAGCTTTGGTATTTCTGC 60
Db 1 GATCAGTGTGTGAGGGAACCTCCATCATGAGGCTGACAAAGTCAGCTTTGGTATTTCTGC 60
QY 61 TCCTGCACTCTTCTGTTGCTGTGATTTCTGTGGAAAGCTCTGTTGGTGGCCCTGTG 120
Db 61 TCCTGCACTCTTCTGTTGCTGTGATTTCTGTGGAAAGCTCTGTTGGTGGCCCTGTG 120
QY 121 ACATGAGCCATTGGCTTAATGTCAGGTCATTTCTAGNAGAGCTCATAGTAGAGGCCATG 180
Db 121 ACATGAGCCATTGGCTTAATGTCAGGTCATTTCTAGNAGAGCTCATAGTAGAGGCCATG 180
QY 181 AGGTAAACAGTATTGACTCACTCAAGGCTTCTGTTAATTAATGACTACAGGAAGCTTTCTGCAT 240
Db 181 AGGTAAACAGTATTGACTCACTCAAGGCTTCTGTTAATTAATGACTACAGGAAGCTTTCTGCAT 240
QY 241 TGAATTTGAGGTGGTCCATATGCCACAGGACAGACAGAAAGTAAATATTTGTTG 300
Db 241 TGAATTTGAGGTGGTCCATATGCCACAGGACAGACAGAAAGTAAATATTTGTTG 300
QY 301 ACCTAGCTCTGAATGCTTTCAGGCTTATCCAGGCTTATCAACCTGGCAATCAGTTATAAATTAATG 360
Db 301 ACCTAGCTCTGAATGCTTTCAGGCTTATCAACCTGGCAATCAGTTATAAATTAATG 360
```



```
QY 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATAATGATGTGTGAGAGCTTTATCTACAATC 420
Db 361 ATTTTGTGTTGAAATAGAGGAACCTTTAAATAATGATGTGTGAGAGCTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCCCTGTGA 480
Db 421 AGACGCTTATGAAGAAGCTACAGGAACCAACTACGATGTAATGCTTATAGACCCCTGTGA 480
QY 481 TTCCCTGTGGAGACCTGATGCTGAGTTGCTGTGAGTCCCTTTTGTGCTCACACTTGA 540
Db 481 TTCCCTGTGGAGACCTGATGCTGAGTTGCTGTGAGTCCCTTTTGTGCTCACACTTGA 540
QY 541 TTTCTGTAGGAGGAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTCTATG 600
Db 541 TTTCTGTAGGAGGAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTCTATG 600
QY 601 TACCTGTCCCTATGACAGGACTTAACAGACAGAATGACCTTTTCTGGAAAGATTAATAAT 660
Db 601 TACCTGTCCCTATGACAGGACTTAACAGACAGAATGACCTTTTCTGGAAAGATTAATAAT 660
QY 661 CAATGCTTTTCACTTTTGTTCACACTTCTGGATTTCAGGATTACGACTATCATTTTGGGA 720
Db 661 CAATGCTTTTCACTTTTGTTCACACTTCTGGATTTCAGGATTACGACTATCATTTTGGGA 720
QY 721 AGTTTATAGTAAGGCATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
Db 721 AGTTTATAGTAAGGCATTAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
QY 781 AGATATGGCTAATACGAACATATGCGGATTTTGAATTTTCCCTAACCATACCACTAACT 840
Db 781 AGATATGGCTAATACGAACATATGCGGATTTTGAATTTTCCCTAACCATACCACTAACT 840
QY 841 TTGAGTTTGTGGAGGATTGCACTTAACTGCGCAAGCTTTGCTAAGGAAATGGAA 900
Db 841 TTGAGTTTGTGGAGGATTGCACTTAACTGCGCAAGCTTTGCTAAGGAAATGGAA 900
QY 901 ATTTTGTCCAGAGTTTCAGGGAAGATGTTATGTTGTTGTTTCTCTGCGGTCACCTGTTTC 960
Db 901 ATTTTGTCCAGAGTTTCAGGGAAGATGTTATGTTGTTTCTCTGCGGTCACCTGTTTC 960
QY 961 AAAATGTTACGAAGAAGGCTTAATCATTTGCTTTCAGCCCTTCCAGCAATCTCGGCTG 1020
Db 961 AAAATGTTACGAAGAAGGCTTAATCATTTGCTTTCAGCCCTTCCAGCAATCTCGGCTG 1020
QY 1021 AGTGTTATGAGGTTACAAAGGAAACCAATCCATTTAGAGCCATCTCGGCTG 1080
Db 1021 AGTGTTATGAGGTTACAAAGGAAACCAATCCATTTAGAGCCATCTCGGCTG 1080
QY 1081 ATGATTGGATACCCAGAAATGATCTTTTGGTTCATCCCAAAACCAAGCTTTTATCCTC 1140
Db 1081 ATGATTGGATACCCAGAAATGATCTTTTGGTTCATCCCAAAACCAAGCTTTTATCCTC 1140
QY 1141 ATGGTGGAAATGAATGGGATCTATGAAGCTATTTACATGCGGTCCTATGTTGGGATTC 1200
Db 1141 ATGGTGGAAATGAATGGGATCTATGAAGCTATTTACATGCGGTCCTATGTTGGGATTC 1200
QY 1201 CCATATTTGGTATCAGCTTGTGATACATAGCTCAGATGAGGCCAAAGGAGCAGCTGTAG 1260
Db 1201 CCATATTTGGTATCAGCTTGTGATACATAGCTCAGATGAGGCCAAAGGAGCAGCTGTAG 1260
QY 1261 AAATAAATCTCAAACTATGACAAGCGAAGATTTTACTGAGGCTTTGAGAAGAGCTCATTA 1320
Db 1261 AAATAAATCTCAAACTATGACAAGCGAAGATTTTACTGAGGCTTTGAGAAGAGCTCATTA 1320
QY 1321 CCGATTTCCTTATAAAGAGATGCTATGAGATTTATCAAGATTTACCATGATCAACCTG 1380
Db 1321 CCGATTTCCTTATAAAGAGATGCTATGAGATTTATCAAGATTTACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTAGATCGAGCAGTCTTCTGATCGAGTTTCTGATCGAGTTTCTGATCGAGTTT 1440
Db 1381 TAAAGCCCTAGATCGAGCAGTCTTCTGATCGAGTTTCTGATCGAGTTTCTGATCGAGTTT 1440
```

```
QY 1441 AGCACCTGCGATCAGCTGCCCATGACCTCAGCTGGTCCAGCAGCTACTCTATAGATGTGA 1500
Db 1441 AGCACCTGCGATCAGCTGCCCATGACCTCAGCTGGTCCAGCAGCTACTCTATAGATGTGA 1500
QY 1501 TTGGGTTCCCTGCTGACCTGTGTGCAACTGCTATATCTTCTCACAAAATGTTTTTAT 1560
Db 1501 TTGGGTTCCCTGCTGACCTGTGTGCAACTGCTATATCTTCTCACAAAATGTTTTTAT 1560
QY 1561 TTTCTGTCAAAAATTTTAAATAAACTAGAAAAGATAGAAAAGAGGGAATPAGATCTTTCCAA 1620
Db 1561 TTTCTGTCAAAAATTTTAAATAAACTAGAAAAGATAGAAAAGAGGGAATPAGATCTTTCCAA 1620
QY 1621 ATTCAAGAAAGACCTG 1636
Db 1621 ATTCAAGAAAGACCTG 1636

RESULT 13
US-10-180-557-521
; Sequence 521, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC147
; CURRENT APPLICATION NUMBER: US/10/180,557
; PRIOR FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-557-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGTGTGTGAGGGAAGTCCCATCATGAGTCTGTGACAAAGTCAGCTTTGGTATTTCTGC 60
Db 1 GATCAGTGTGTGAGGGAAGTCCCATCATGAGTCTGTGACAAAGTCAGCTTTGGTATTTCTGC 60
QY 61 TCCGTGAGCTCTTCTGTGTGTGCTGTGGATTCTGTGGGAAAGTCTCTGGTGTGGCCCTGTG 120
Db 61 TCCGTGAGCTCTTCTGTGTGTGCTGTGGATTCTGTGGGAAAGTCTCTGGTGTGGCCCTGTG 120
QY 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCATTTCTAGAAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTTGGCTTAAATGTCAAGGTCATTTCTAGAAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAACAGTATTGACTCACTCAAAAGCTTTCGTTAATTGACTACAGGAAGCCCTTCGCAAT 240
Db 181 AGGTAACAGTATTGACTCACTCAAAAGCTTTCGTTAATTGACTACAGGAAGCCCTTCGCAAT 240
QY 241 TGAATTTGAGGTGTGCTCATATGCCACAGCAGACAGACAGAAAATGAAATATTTGTTG 300
Db 241 TGAATTTGAGGTGTGCTCATATGCCACAGCAGACAGACAGAAAATGAAATATTTGTTG 300
QY 301 ACCTAGCTCTGAATGTCTTCCAGGCTTATCAACCTGCGCAATCAGTTATAAATAATTAATG 360
Db 301 ACCTAGCTCTGAATGTCTTCCAGGCTTATCAACCTGCGCAATCAGTTATAAATAATTAATG 360
```

Qy 361 ATTTTGTGTTGAATAGAGAACTTTAAAAATGATGTGTGAGAGCTTTTATCTACAATC 420
Db 361 ATTTTGTGTTGAATAGAGAACTTTAAAAATGATGTGTGAGAGCTTTTATCTACAATC 420
Qy 421 AGAGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTATAGACCCTGTGA 480
Db 421 AGAGCTTATGAAGAAGCTACAGAAACCAACTACGATGTAATGCTTATAGACCCTGTGA 480
Qy 481 TTCCCTGTGGAGCTGATGCTCAGTGTGCTTCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGGAGCTGATGCTCAGTGTGCTTCAGTCCCTTTTGTGCTCACACTTAGAA 540
Qy 541 TTTCTGTAGGAGCAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCCTATG 600
Db 541 TTTCTGTAGGAGCAATATGAGGGAAGCTGTGGGAAACTTCCAGCTCCACTTTTCCCTATG 600
Qy 601 TACCTGTGCTTATGACAGGACTACACAGACAATGACCTTTCTGGAAGAGTAATAAATT 660
Db 601 TACCTGTGCTTATGACAGGACTACACAGACAATGACCTTTCTGGAAGAGTAATAAATT 660
Qy 661 CAATGCTTTTCACTTTTGTCCACTTCTGGATTGAGGATTACGACTATATCATTTTGGGAAG 720
Db 661 CAATGCTTTTCACTTTTGTCCACTTCTGGATTGAGGATTACGACTATATCATTTTGGGAAG 720
Qy 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
Db 721 AGTTTATAGTAAGGCAATAGGAAGGCCACTACATTTATGTGAGACTGTGGGAAAGCTG 780
Qy 781 AGATATGCTAATACGAACATATTTGGGATTTTGAATTTTCTTCAACCATACCACTAACT 840
Db 781 AGATATGCTAATACGAACATATTTGGGATTTTGAATTTTCTTCAACCATACCACTAACT 840
Qy 841 TTGAGTTTGTGGAGGATTGACATGTAACCTGCAAGCTTTCGCTAAGCAATAGGAA 900
Db 841 TTGAGTTTGTGGAGGATTGACATGTAACCTGCAAGCTTTCGCTAAGCAATAGGAA 900
Qy 901 ATTTTGTCCAGAGTTTCAGGGGAAGATGTTATGTTGGTGTTCCTCTGGGGTCACTGTTTC 960
Db 901 ATTTTGTCCAGAGTTTCAGGGGAAGATGTTATGTTGGTGTTCCTCTGGGGTCACTGTTTC 960
Qy 961 AAAATGTTACAGAAAGGCTAATATCATTTGCTTACAGCCCTTGGCCAGATCCACAGA 1020
Db 961 AAAATGTTACAGAAAGGCTAATATCATTTGCTTACAGCCCTTGGCCAGATCCACAGA 1020
Qy 1021 AGGTGTTATGAGGTACAAAGGAAACCAATCCATAGGAGGCAATACCTCGGCTGT 1080
Db 1021 AGGTGTTATGAGGTACAAAGGAAACCAATCCATAGGAGGCAATACCTCGGCTGT 1080
Qy 1081 ATGATTTGGATACCCAGAAATGATCTTCTTGTGTCATCCCAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTTGGATACCCAGAAATGATCTTCTTGTGTCATCCCAACCAAGCTTTTATCACTC 1140
Qy 1141 ATGTTGGAATGAATGGGATCTATGAAGCTATTTTACCATGGGGTCCCTATGTTGGGAGTTC 1200
Db 1141 ATGTTGGAATGAATGGGATCTATGAAGCTATTTTACCATGGGGTCCCTATGTTGGGAGTTC 1200
Qy 1201 CCATATTTGGTGATCAGCTTCATTAACATAGCTACATGAGGCGCAAGGAGCAGCTGTAG 1260
Db 1201 CCATATTTGGTGATCAGCTTCATTAACATAGCTACATGAGGCGCAAGGAGCAGCTGTAG 1260
Qy 1261 AAATTAACCTTCAAACTATGACAGCGAAGATTTTACTGAGGGCTTTTGAAGACAGTCAPTA 1320
Db 1261 AAATTAACCTTCAAACTATGACAGCGAAGATTTTACTGAGGGCTTTTGAAGACAGTCAPTA 1320
Qy 1321 CCGATTTCTTATTAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Db 1321 CCGATTTCTTATTAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Qy 1381 TAAAGCCCTAGATCGAGCAGCTTCTGAGTTCAGTGTGCTATGCGCCACAAAGAGCCA 1440
Db 1381 TAAAGCCCTAGATCGAGCAGCTTCTGAGTTCAGTGTGCTATGCGCCACAAAGAGCCA 1440

RESULT 14

US-10-173-700-521
; Sequence 521, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173.700
; CURRENT FILING DATE: 2002-05-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-700-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGCAAGTGCAGCTTTGGTATTTCTGC 60
Db 1 GATCAGTGTGTGAGGAACTGCCATCATGAGGTCTGCAAGTGCAGCTTTGGTATTTCTGC 60
Qy 61 TCCTGCAGCTCTTCTGTGTTGGCTGTGGATTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG 120
Db 61 TCCTGCAGCTCTTCTGTGTTGGCTGTGGATTCTGTGGAAAGTCTCTGGTGTGGCCCTGTG 120
Qy 121 ACATGAGCCATTTGCTTAATGTCAAGTTCATCTTAGAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTTGCTTAATGTCAAGTTCATCTTAGAGAGCTCATAGTGAGAGGCCATG 180
Qy 181 AGGTAACAGTATTGACCTCAAGGCTTCGTTAATTTGACTACAGGAAGCTTTCTGCAT 240
Db 181 AGGTAACAGTATTGACCTCAAGGCTTCGTTAATTTGACTACAGGAAGCTTTCTGCAT 240
Qy 241 TGAATTTGAGGTGGTCCATATGCCAGGACAGACAGAAATGAAATATTTGTTG 300
Db 241 TGAATTTGAGGTGGTCCATATGCCAGGACAGACAGAAATGAAATATTTGTTG 300
Qy 301 ACCTAGCTCTGAATGCTCTTCCAGGCTTATCAACCTGGCATCAGTTATAAATTAATG 360
Db 301 ACCTAGCTCTGAATGCTCTTCCAGGCTTATCAACCTGGCATCAGTTATAAATTAATG 360

QY 361 ATTTTGTGTTGTAATGAGGAGCACTTTAAATAATGATGTTGAGAGCTTTATCTACAATC 420
Db 361 ATTTTGTGTTGTAATGAGGAGCACTTTAAATAATGATGTTGAGAGCTTTATCTACAATC 420
QY 421 AGACGCTTATGAAGAAGCTACAGGAACCACTACGATGTAATGCTTATAGACCCGTGCA 480
Db 421 AGACGCTTATGAAGAAGCTACAGGAACCACTACGATGTAATGCTTATAGACCCGTGCA 480
QY 481 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
Db 481 TTCCCTGTGAGACCTGATGGCTGAGTGTGCTTGCAGTCCCTTTTGTGCTCACACTTAGAA 540
QY 541 TTCTCTGTAGAGGCAATATGAGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 600
Db 541 TTCTCTGTAGAGGCAATATGAGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 600
QY 601 TACCTGTGCTATGACAGGACTAACACAGAGATGACCTTTCTGGAAGAGTAAAAAAT 660
Db 601 TACCTGTGCTATGACAGGACTAACACAGAGATGACCTTTCTGGAAGAGTAAAAAAT 660
QY 661 CAATGCTTTTCCAGTTTGTCCACTTCTGGATTCCAGGATACGACTATCATTTTGGGAAG 720
Db 661 CAATGCTTTTCCAGTTTGTCCACTTCTGGATTCCAGGATACGACTATCATTTTGGGAAG 720
QY 721 AGTTTATAGTAGGCAATATGAGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 780
Db 721 AGTTTATAGTAGGCAATATGAGGAGCAAGCTGTGGGAAACTTCCAGCTCCACTTTTCTATG 780
QY 781 AGATATGGCTTAATACGAGCAATATGAGGATTTTGAATTTTCTCAACCATACCAACCTAACT 840
Db 781 AGATATGGCTTAATACGAGCAATATGAGGATTTTGAATTTTCTCAACCATACCAACCTAACT 840
QY 841 TTGAGTTTGTGGAGGATGACGTTAAACCTGCGCAAGCTTTGCCCTAAGGAATGAAA 900
Db 841 TTGAGTTTGTGGAGGATGACGTTAAACCTGCGCAAGCTTTGCCCTAAGGAATGAAA 900
QY 901 ATTTTGTCCAGGTTTCCAGGAGAGATGATGTTGGTGTCTTCTCGGGTCACTGTTTC 960
Db 901 ATTTTGTCCAGGTTTCCAGGAGAGATGATGTTGGTGTCTTCTCGGGTCACTGTTTC 960
QY 961 AAAATGTTACAGAGAAAGCTTAATATCATTTGCTTCCAGCCCTTCCAGATCCACAGA 1020
Db 961 AAAATGTTACAGAGAAAGCTTAATATCATTTGCTTCCAGCCCTTCCAGATCCACAGA 1020
QY 1021 AGGTGTTATGAGGATACAAAGGAAAAAACCATTCCAGATTAGGAGCAATACCTCGCTGT 1080
Db 1021 AGGTGTTATGAGGATACAAAGGAAAAAACCATTCCAGATTAGGAGCAATACCTCGCTGT 1080
QY 1081 ATGATTTGGATACCCAGAAATGATCTTCTTGTGTCATCCCAAAACCAAGCTTTTATCACTC 1140
Db 1081 ATGATTTGGATACCCAGAAATGATCTTCTTGTGTCATCCCAAAACCAAGCTTTTATCACTC 1140
QY 1141 ATGTTGGAATGAATGGGATCATGAAGCTATTTACCATGGGGTCCCTATGTTGGAGTTC 1200
Db 1141 ATGTTGGAATGAATGGGATCATGAAGCTATTTACCATGGGGTCCCTATGTTGGAGTTC 1200
QY 1201 CCATATTTGGTATCAGCTTGTATACATAGCTACATGAAGCCCAAGGAGCAGCTGTAG 1260
Db 1201 CCATATTTGGTATCAGCTTGTATACATAGCTACATGAAGCCCAAGGAGCAGCTGTAG 1260
QY 1261 AAATAAATCTCAAAACTATGACAGCGAAGATTTTACTGAGGCTTTTGAGAACAGTCATTA 1320
Db 1261 AAATAAATCTCAAAACTATGACAGCGAAGATTTTACTGAGGCTTTTGAGAACAGTCATTA 1320
QY 1321 CCGATTTCCTTATTAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
Db 1321 CCGATTTCCTTATTAAGAGAAATGCTATGAGATTTATCAAGAAATTCACCATGATCAACCTG 1380
QY 1381 TAAAGCCCTTACATCGACGCTTCTGATCGAGTGTGTCATGCGCCACAAGAGCA 1440
Db 1381 TAAAGCCCTTACATCGACGCTTCTGATCGAGTGTGTCATGCGCCACAAGAGCA 1440

QY 1441 AGCACTCGCATCAGCTGCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGA 1500
Db 1441 AGCACTCGCATCAGCTGCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGA 1500
QY 1501 TTGGGTTCTGCTGACCTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTTAT 1560
Db 1501 TTGGGTTCTGCTGACCTGTGGCAACTGCTATATCTTGTTCACAAAATGTTTTTAT 1560
QY 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAGATGAAAAGAGGGAATAGATCTTTCCAA 1620
Db 1561 TTTCCTGTCAAAAATTTAATAAACTAGAAAGATGAAAAGAGGGAATAGATCTTTCCAA 1620
QY 1621 ATTCAAGAAAGACCTG 1636
Db 1621 ATTCAAGAAAGACCTG 1636

RESULT 15
US-10-174-572-521
; Sequence 521, Application US/10174572
; Publication No. US20030027263A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C40
; CURRENT APPLICATION NUMBER: US/10/174, 572
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 521
; LENGTH: 2974
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-572-521

Query Match 100.0%; Score 1636; DB 9; Length 2974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGCAAGTCTGAGCTTTGGTATTTCTGC 60
Db 1 GATCAGTGTGTGAGGGAACCTGCCATCATGAGGTCTGCAAGTCTGAGCTTTGGTATTTCTGC 60
QY 61 TCCTGCAGCTTCTGTTGCTGTGGGAAAGTCTGTTGGGAAAGTCTGTTGGGCTGTG 120
Db 61 TCCTGCAGCTTCTGTTGCTGTGGGAAAGTCTGTTGGGAAAGTCTGTTGGGCTGTG 120
QY 121 ACATGAGCCATTGGCTTAATGTCAAGTCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
Db 121 ACATGAGCCATTGGCTTAATGTCAAGTCTTCTAGAGAGCTCATAGTGAGAGGCCATG 180
QY 181 AGGTAAACAGTATTGACTCAGTCAAGCCCTTCGTTAAATTTGACTACAGGAGCCCTCTGCAT 240
Db 181 AGGTAAACAGTATTGACTCAGTCAAGCCCTTCGTTAAATTTGACTACAGGAGCCCTCTGCAT 240
QY 241 TGAATTTTGTGAGTGTCCATATGCCACAGGAGCAGACAGAAAATGAAATATTTTGTG 300
Db 241 TGAATTTTGTGAGTGTCCATATGCCACAGGAGCAGACAGAAAATGAAATATTTTGTG 300
QY 301 ACCTAGCTCTGAATGCTTGTCCAGGCTTATCAACCTGGCAATCAGTGTATATAAATG 360
Db 301 ACCTAGCTCTGAATGCTTGTCCAGGCTTATCAACCTGGCAATCAGTGTATATAAATG 360

Qy	361	ATTTT	TTGTTGAAAT	TAAGAGGAAC	TTTAAAAA	TGATGTGTGAGAC	TTTATCTACAATC	420	
Db	361	ATTTT	TTGTTGAAAT	TAAGAGGAAC	TTTAAAAA	TGATGTGTGAGAGC	TTTATCTACAATC	420	
Qy	421	AGACG	CTTATGAGAAG	CTACAGAAAC	CAACCTAC	GATGTAAT	TGCTTATAGACCC	480	
Db	421	AGACG	CTTATGAGAAG	CTACAGAAAC	CAACCTAC	GATGTAAT	TGCTTATAGACCC	480	
Qy	481	TTCCCT	GTGAGAGAC	CTGATGGCT	TGAGTTGCT	TGCAGTCC	CTTTTGTGCTCACACT	540	
Db	481	TTCCCT	GTGAGAGAC	CTGATGGCT	TGAGTTGCT	TGCAGTCC	CTTTTGTGCTCACACT	540	
Qy	541	TTTCTG	TAGAGGCA	ATATGGAG	GGAAGCTGTGGG	AAACCTTCC	CAGCTCCACTTTTCTCTATG	600	
Db	541	TTTCTG	TAGAGGCA	ATATGGAG	GGAAGCTGTGGG	AAACCTTCC	CAGCTCCACTTTTCTCTATG	600	
Qy	601	TACCTG	TGCCCTATG	CACAGG	ACTAAC	CACAGAGATG	ACCTTTCTTGGG	AAAGAGTAAAAA	660
Db	601	TACCTG	TGCCCTATG	CACAGG	ACTAAC	CACAGAGATG	ACCTTTCTTGGG	AAAGAGTAAAAA	660
Qy	661	CAATG	CTTTCAG	TTTTTGTTC	CCACTCTCG	GATTCAGG	ATTACGACTATCAT	720	
Db	661	CAATG	CTTTCAG	TTTTTGTTC	CCACTCTCG	GATTCAGG	ATTACGACTATCAT	720	
Qy	721	AGTTT	TATAGTAAG	GCATTAAG	GAGCCCACTACA	TTATGTG	AGAGCTGTGGG	AAAAAGCTG	780
Db	721	AGTTT	TATAGTAAG	GCATTAAG	GAGCCCACTACA	TTATGTG	AGAGCTGTGGG	AAAAAGCTG	780
Qy	781	AGATG	GCTTAATAC	GNACATAT	TGGATTTT	TGAAATTT	CCTCAACCATACCA	ACCTTAAC	840
Db	781	AGATG	GCTTAATAC	GNACATAT	TGGATTTT	TGAAATTT	CCTCAACCATACCA	ACCTTAAC	840
Qy	841	TTGAG	TTTGTGG	AGGATTGC	ACTGTAA	AGCTTGC	CAAGCTTTTGCCT	TAAGGAAATGGAA	900
Db	841	TTGAG	TTTGTGG	AGGATTGC	ACTGTAA	AGCTTGC	CAAGCTTTTGCCT	TAAGGAAATGGAA	900
Qy	901	ATTTT	TGTCAGAG	TTCAGGGG	AGATGG	TATGTGTG	TGTTTTCTCTG	GGGTCACTGTTTC	960
Db	901	ATTTT	TGTCAGAG	TTCAGGGG	AGATGG	TATGTGTG	TGTTTTCTCTG	GGGTCACTGTTTC	960
Qy	961	AAATG	TTCAGAA	GAAGGCTTA	TATCA	TTCTC	TGAGCTTTGCC	AGATCCACAGA	1020
Db	961	AAATG	TTCAGAA	GAAGGCTTA	TATCA	TTCTC	TGAGCTTTGCC	AGATCCACAGA	1020
Qy	1021	AGGTG	TATGGAG	GTACAA	AGGAAAAA	AAACCATCC	ACATTAGGAG	CCAACTAC	1080
Db	1021	AGGTG	TATGGAG	GTACAA	AGGAAAAA	AAACCATCC	ACATTAGGAG	CCAACTAC	1080
Qy	1081	ATGAT	TGGATAC	CCCGAG	ATGATCT	TCTGGT	CATCCCA	AAACCAAGCTTTTAT	1140
Db	1081	ATGAT	TGGATAC	CCCGAG	ATGATCT	TCTGGT	CATCCCA	AAACCAAGCTTTTAT	1140
Qy	1141	ATGGT	GGAATGA	ATGGG	ATCATGA	AGCTATTT	TACCATGGG	GTCCCTATGGTGGG	1200
Db	1141	ATGGT	GGAATGA	ATGGG	ATCATGA	AGCTATTT	TACCATGGG	GTCCCTATGGTGGG	1200

Search completed: July 16, 2003, 18:04:24
Job time : 220.273 secs

Qy	1441	AGCACCTGCGATCAGTGGCCCATGACCTCACTGGTTCCAGCACTACTCTATAGATGTGA	1500
Db	1441	AGCACCTGCGATCAGTGGCCCATGACCTCACTGGTTCCAGCACTACTCTATAGATGTGA	1500
Qy	1501	TTGGGTTCTGCTGCAGCTGTGGCAACTGCTATATTCTTCACAAAAATGTTTTTTTAT	1560
Db	1501	TTGGGTTCTGCTGCAGCTGTGGCAACTGCTATATTCTTCACAAAAATGTTTTTTTAT	1560
Qy	1561	TTTCTCTGTCAAAAATTTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTTCCAA	1620
Db	1561	TTTCTCTGTCAAAAATTTTAATAAACTAGAAAAGATAGAAAAGAGGGAATAGATCTTTTCCAA	1620
Qy	1621	ATTCAAGAAAGACCTG	1636
Db	1621	ATTCAAGAAAGACCTG	1636